CNA

October 2021

The Psychology of (Dis)information: Case Studies and Implications

Megan McBride, Heather Wolters, Kaia Haney, and William Rosenau

With contributions by Neil Carey and Kasey Stricklin

DRM-2021-U-030881-Final

Abstract

The absorption and spread of disinformation is a pervasive phenomenon across a wide variety of topics and media. Most disinformation research focuses on the source (who created it?) and the environment in which it exists (what platform/medium transmits the information?). Recognizing that disinformation primarily works in an individual person's mind, this report describes four normal, routine psychological mechanisms that are associated with the absorption and spread of disinformation. We then describe real-world case studies—focusing on activities linked to COVID-19, and to campaigns coordinated by US adversaries including Russia, China, and Iran—to illustrate the way these mechanisms can be manipulated to aid the spread disinformation. The report concludes with multi-pronged recommendations that DOD can use to address the vulnerabilities associated with these psychological mechanisms so as prevent the spread of disinformation and protect both US servicemembers and the country.

Distribution

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

10/1/2021

This work was performed under Federal Government Contract No. N00014-16-D-5003.

Cover image credit: Adapted from Cristina Spanò; "Information Overload Helps Fake News Spread, and Social Media Knows it," Scientific American [Online], Dec. 1, 2020, accessed Sept. 27, 2021, https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/.

Approved by: **October 2021**

Jonathan Schroden, Research Program Director
Countering Threats and Challenges Program
Strategy, Policy, Plans, and Programs Division

Request additional copies of this document through inquiries@cna.org.

Executive Summary

The absorption and spread of disinformation[1] is a pervasive phenomenon across a wide variety of topics on virtually every social media network. The avalanche of COVID-19 disinformation that has been produced over the last 18 months typifies the prevalence of disinformation in the modern world. In fact, a 2018 MIT study of Twitter data found that disinformation was more successful than truth on social media by almost every known metric; it spread "significantly farther, faster, deeper, and more broadly than the truth in all categories of information."[2]

Further, the absorption and spread of disinformation is a growing national security concern. Most of our adversaries recognize that controlling the information space, including domestic and international narratives, promotes their national interests. For example, in 2013, the Russian chief of the General Staff, Valeriy Gerasimov, stated that the development of information weapons had the ability to reduce an adversary's combat potential.[3] Also, in 2013, Chinese president Xi Jinping stated that the use of innovative techniques to spread narratives positive for China, and promoting the Chinese view globally, was a priority.[4]

Although the use of information operations is not a new phenomenon—various actors have used them throughout history for a range of objectives—the connectivity that characterizes the world today allows both information and disinformation to spread faster and with a much greater reach. The use of disinformation has led directly to real-world events and violence. It can have a demonstrable effect on a recipient's behavior, and can lead its promulgators to achieve some goals simply through its existence, regardless of its believability.

Recognizing that disinformation's primary effect is on the mind, this report describes four psychological mechanisms that are associated with the absorption and spread of disinformation. It then connects five, recent, real-world examples of disinformation in which absorption and spread benefited from the psychological mechanisms described. Finally, it

[1] *Disinformation* is the intentional creation and spread of false information.

[2] Soroush Vosoughi, Deb Roy, and Sinan Aral, "The spread of true and false news online," *Science* 359, no. 6380 (2018), https://science.sciencemag.org/content/359/6380/1146.

[3] V. V. Gerasimov, "The Value of Science is Foresight," Ценность науки в предвидении, VPK, ВПК, Feb. 26, 2013, https://vpk-news.ru/articles/14632.

[4] Daniel Kliman et al., *Dangerous Synergies*, CNAS, 2020, https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/CNAS-Report-Dangerous-Synergies-May-2020-DoS-Proof.pdf?mtime=20200506164642&focal=none.

describes possible ways the Department of Defense (DOD) can mitigate the absorption and spread of disinformation among US servicemembers by focusing on the psychological mechanisms that contribute to its promulgation.

Psychological mechanisms associated with disinformation spread

This report describes four key psychological mechanisms and explains how each one contributes to the absorption and spread of disinformation:[5]

- **Initial information processing:** Our mental "processing capacity" is limited; we simply cannot deeply attend to all new information we encounter. Our brains take mental shortcuts to incorporating new information, and those shortcuts can open us up to mistakes. To the extent that we do not process information as thoroughly as we should, we can construe disinformation as true information.
- **Cognitive dissonance:** *Cognitive dissonance* describes the discomfort we feel when we are confronted with two competing ideas and wish to reduce that discomfort. If disinformation supports our initial beliefs or creates less dissonance than true information, we are more likely to believe the disinformation.
- **Influence of group membership, beliefs, and novelty (the GBN model):** Not all information is equally valuable to individuals. Our group memberships, our beliefs, and the uniqueness of the information influence whether we absorb and share disinformation. We are more likely to share information with people we consider members of our group, when we believe the information is true, and when it is novel or urgent. If disinformation comes from a group member with whom we identify, is consistent with our beliefs, or is new information for us, we are more likely to share it.
- **Role of emotion and arousal in our sharing of disinformation:** Just as not all information is equally valuable, not all information affects us the same way. Research demonstrates that we pay more attention to information that makes us feel positively (i.e., good) or that arouses us to act. Given that disinformation is, by definition, created by someone, it is more likely to be absorbed and shared if it is constructed to be emotional and arousing.

A critical takeaway from the identification of these mechanisms is that they are not unique to absorbing and spreading disinformation. These same mechanisms are key to absorbing and

[5] This literature review is documented in the companion to this report: Heather Wolters et al., *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*, CNA, DRM-2021-U-029337-1Rev, Sept. 2021.

spreading true information as well. Thus, at an individual level, it appears that disinformation is absorbed and spread through normal, routine, and adaptive mechanisms, which malign actors can exploit and manipulate for their own objectives.

Real-world case studies help illuminate the psychological mechanisms

To illustrate the psychological mechanisms involved with disinformation, we researched recent real-world examples. We selected five case studies, included in this report, because they are examples of disinformation campaigns executed by US adversaries. In aggregate, this portfolio of cases addresses a wide variety of domestic and international issues affecting US interests both at home and abroad; they illustrate the full range of psychological mechanisms that this report looks to address. Although there were multiple psychological mechanisms associated with each case study, we focused our analysis and discussion on the most salient in order to clearly describe the primary mechanisms operating in each example.

Operation Secondary Infektion was a complex, multi-year effort, beginning in 2014, by an unknown Russian entity to plant fake news, forged documents, and divisive content on multiple social-media platforms (including Google, Facebook, Reddit, Twitter, and YouTube). The primary objective was to foment discord and division within countries and institutions Russia perceived as its adversaries. This case illustrates the psychological mechanisms of (1) *initial information processing,* through the use of forged documents and the mainstream media, (2) *the GBN model,* through urgent and novel information about a possible assassination, and (3) *the role of emotions and arousal* through multiple mentions to known terrorist groups and inflammatory political issues.

Endless Mayfly has been active since at least 2016 and is an Iran-linked disinformation campaign that includes a network of false personas, social media accounts, and websites that impersonate legitimate news organizations. This case illustrates the psychological mechanisms of *initial information processing*. It meticulously spoofed legitimate websites and organizations such that, if someone was reading quickly and not paying careful attention, they might perceive the source as legitimate.

During the **2016 US presidential election**, the FBI began to investigate an extensive Russian operation aimed at influencing the election by manipulating Americans and causing social divisions through social media platforms. This case illustrates the psychological mechanisms of (1) *cognitive dissonance,* through customizing information that conforms to pre-existing beliefs, thereby reducing scrutiny and increasing the individual's commitment to those beliefs; and (2) *the role of emotions and arousal,* through the creation of divisive content on

controversial issues regarding race, patriotism, immigration, gun control, and LGBT rights, with the goal of stoking fear, anger, and division among Americans.

Hong Kong: Protestors or Terrorists refers to a 2019 disinformation campaign in which the Chinese government sought to portray the protests in Hong Kong as terrorist acts. This influence campaign was global, targeting Chinese citizens, members of the Chinese diaspora internationally, Hong Kong residents, and the international community. This case illustrates the psychological mechanisms of (1) *cognitive dissonance,* through the creation of a single and consistent narrative that minimizes dissonance for the recipient and any need to mitigate discomfort from that dissonance; and (2) the *role of emotions and arousal,* through the creation and dissemination of imagery meant to stimulate patriotic feelings, fear, or disgust.

Finally, **#CoronaJihad** was a disinformation campaign that began in March 2020, just as the COVID-19 pandemic was starting. The campaign targeted an Islamic reformist group in India by distorting and falsifying information related to the group's activities at the beginning of the pandemic. This case illustrates the psychological mechanisms of (1) *the GBN model,* by setting as rivals groups that have long-standing tensions between them, and (2) *the role of emotions and arousal,* by focusing on an imminent threat and evoking feelings of fear and disgust.

These five real-world examples of disinformation clearly illustrate the human vulnerability to what we might call *psychological hacking.* Each of them demonstrates how the manipulation of normal and functional psychological mechanisms can increase our receptivity to disinformation, and how those who create disinformation can leverage this vulnerability by designing content that is more likely to evade scrutiny or increase engagement. The cases also illustrate some potentially real—not merely notional—threats that the US faces by describing disinformation campaigns that have already been executed. In doing so, they provide detail on the capabilities that our foreign adversaries already possess, and on the threat posed by grassroots campaigns that go unchecked.

Recommendations for DOD

The absorption and spread of disinformation is very difficult to extinguish. It happens incredibly quickly, through multiple media and social media channels, around the world, and, importantly, in the minds of each individual interacting with the material. As a result, some solutions lie outside the DOD's influence, including regulation of social media, increased technology security and scrutiny on unverified content, and leveraging of sanctions for knowingly creating and spreading disinformation. However, DOD is already working on technological means of thwarting state and nonstate actors spreading disinformation in and about the US, and an increasingly robust conversation is already occurring about legislative

action that might force the more aggressive removal of disinformation from social media platforms.

It is critical, though, that DOD invest in the development and deployment of non-partisan, evidence-based interventions that will protect US servicemembers against this content. As this report makes clear, disinformation has already been used to influence perceptions of allies and adversaries; increase distrust of neighbors and friends; distort understanding of current events; affirm that unflattering stereotypes are valid; and leverage worries and fears to prompt action. Deployed expertly, disinformation can influence US partnerships, increase domestic discord and violence, cripple the US economy, and endanger the health and well-being of the US population. Thus, protecting the health of servicemembers—in this case by protecting them from psychological manipulation—protects those tasked to protect the country, and protects the country from servicemembers who might otherwise unwittingly be vectors of foreign adversary disinformation.

In our literature search and discussions with subject matter experts, the research team identified several techniques that could be useful for countering the absorption and spread of disinformation. The most promising techniques can be grouped into two categories: (1) using preventive inoculation (i.e., warning people about the effects of disinformation and how to spot it); and (2) encouraging deeper, analytic thinking. These two techniques can be woven into training (game-based or more traditional training akin to other mandatory DOD training, such as OPSEC) and awareness campaigns, including technology-based awareness, such as an additional verification step required to share content from social media on DOD computers. These two avenues of intervention are likely to be simple, fast, low cost, and already familiar to servicemembers who engage in training on multiple cyber-related topics and who are exposed to multiple awareness campaigns annually.

More research is needed to determine the effectiveness of any of the recommendations contained in this report. However, these suggestions provide a multi-pronged starting point for addressing the psychological mechanisms associated with the absorption and spread of disinformation.

This page intentionally left blank.

Contents

This page intentionally left blank.

Introduction

Disinformation, on a wide variety of topics, is on every social media network; if you are online, you have very likely encountered disinformation. In fact, a 2020 analysis by NewsGuard found that the percentage of unreliable content people were encountering on social media platforms had more than doubled from the previous year. Specifically, they reported that in 2019, 8 percent of "engagement among the top 100 news sources on social media came from sources that NewsGuard deems generally unreliable; by 2020, the number was 17 percent."[6]

Understanding the challenge posed by disinformation or untrustworthy content requires recognizing that there are three primary actors motivating this dynamic. First, malicious actors supply this content in pursuit of their aims (e.g., manipulation, monetary gain). Second, social media companies ignore disinformation because regulation might reduce profit.[7] And third, individual users are psychologically vulnerable to engaging with, and passing along, disinformation designed to be appealing. Unfortunately, the design of the social media platforms and the human vulnerability to spreading disinformation together contribute to a system that unintentionally favors the spread of disinformation.

This needn't be the case, though, and this report explores the third factor in this dynamic: psychological vulnerability to disinformation. Although much has been written about the need to improve internet literacy and critical thinking skills,[8] we have focused on the root cause of the problem these interventions are designed to address. Successful disinformation is effectively a form of psychological hacking, and this report shows precisely *how* this hacking occurs. In doing so, it lays out our vulnerabilities—perfectly normal and functional cognitive mechanisms that are being exploited by malicious actors—and makes the case for additional research into the types of interventions most likely to protect individuals from being manipulated.

---

[6] Sara Fischer, "'Unreliable' news sources got more traction in 2020," Axios, Dec. 22, 2020, https://www.axios.com/unreliable-news-sources-social-media-engagement-297bf046-c1b0-4e69-9875-05443b1dca73.html.

[7] Tatyana Hopkins, "Social media companies profiting from misinformation," GWToday, June 19, 2020, https://gwtoday.gwu.edu/social-media-companies-profiting-misinformation.

[8] Darrell M. West, *How to Combat Fake News and Disinformation*, Brookings Institution, 2017, https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation/; Alice Huguet et al., *Exploring Media Literacy Education as a Tool for Mitigating Truth Decay*, RAND, RR-3050-RC, 2019, doi: https://doi.org/10.7249/RR3050; Kristin M. Lord and Katya Vogt, "Strengthen Media Literacy to Win the Fight Against Misinformation," *Stanford Social Innovation Review* (2021), accessed Aug. 18, 2021, https://ssir.org/articles/entry/strengthen_media_literacy_to_win_the_fight_against_misinformation.

More pointedly, its conclusion argues that the DOD should act aggressively to protect US servicemembers from this type of psychological manipulation. The DOD has a long history of both protecting those that serve the nation, and protecting the nation from adversaries that would use servicemembers as a vector for harming our country. Disinformation is not new, but at its current scale it represents a novel threat. Responding to this threat by stopping producers will almost certainly be an endless game of whack-a-mole as producers evolve in response to DOD interventions. And responding to this threat via regulation is a complex legislative process that would take years to resolve. In the meantime, the DOD can act aggressively and quickly to protect both US servicemembers and the nation by designing and implementing a non-partisan, evidence-based set of interventions that will protect against this innate human vulnerability. The goal of this effort, critically, is not to shape opinion, but to prevent opinion from being shaped by inaccurate data.

Key concepts

To make this complex topic accessible and useful for practical application, this study has several important elements. First, for the purposes of this paper, *we make no distinction between disinformation and misinformation*. The two concepts are distinct, with the key difference being the intent of the information's creator (i.e., malicious versus benign); however, this paper focuses on the information itself and the effect on the person viewing or hearing it. In most cases, the recipient does not know who created the information (e.g., the meme, the video, or the email), so the creator's intent is irrelevant to the psychological effect of the information on the audience. We will discuss and include examples of both disinformation and misinformation, categorizing them both as disinformation for simplicity (despite their definitional differences, which we recognize).

Second, the term *disinformation* has come to mean many different things to different people; at various times, a number of distinct, though related, concepts have been classified as disinformation. In this report, *we focus on the elements of disinformation included in a recent State Department funded report*, and we make the conscious decision to leave out some of the other concepts occasionally included in this category. First, we exclude such techniques as the creation of fake accounts (bots, trolls, etc.) and fake communities. For our purposes, we consider these tools and actions to be part of "influence operations," a category that includes disinformation but is not exclusively composed of the disinformation itself. Though disinformation is a type of influence operation, not all influence operations are disinformation, so we also exclude influence operations that do not include disinformation. In other words, this study centers on the psychology of disinformation, and not on the broader psychology of influence operations. By scoping our report in this way, we are able to speak directly to a pressing national security risk in a clear, concise, and focused manner.

In addition, we are approaching this topic from a *nontechnical perspective*, with the intent of gearing explanations toward those with no prior knowledge of this subject. Therefore, we do not include the universe of potentially relevant psychological principles, which that could get overwhelming and unwieldy. The psychological principles discussed in this report emerge from the literature as highly relevant to absorbing and spreading disinformation. Additional information on these principles can be gleaned from the cited material. Moreover, this report provides clear explanations and examples in order to be useful to its intended audience: policy- and decision-makers.

Research questions and methodology

This report—a CNA initiated study—is the most recent in a series of reports focusing on the crucial nexus of technology, information influence, and the effects on national security.[9] Like the others, though, it is not written for those already fluent in the technological challenges being presented. Instead, this series is written for the non-expert, using clear and forthright language, and it articulates non-technical solutions. This is done for several reasons: to draw attention to the fact that not all technological challenges need technological solutions; to acknowledge that those well versed in this technology are already working on technological solutions; and to fill a critical gap in the national security discourse. This particular study aims to answer three key questions:

- Why do people engage with disinformation? Why are they persuaded by the content of disinformation? What are the relevant psychological mechanisms that explain the effectiveness of disinformation?
- How do adversaries employ these psychological mechanisms in the construction of disinformation?
- What are the implications of these findings for US government and DOD actors?

In order to answer these questions, our research team first conducted a comprehensive interdisciplinary literature review capturing relevant information from fields including disinformation, psychology, military science, foreign affairs, economics, computer science, and marketing. These data were used to produce a companion to this report—*The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*—that explains, in detailed but

[9] Vera Zakem, Megan K. McBride, and Kate Hammerberg, *Exploring the Utility of Memes for U.S. Government Influence Campaigns*, CNA, 2018, accessed Aug. 17, 2021, https://shb2016docweb.cna.org:8443/dctmsearch/FFRDC/Publications/2018/DRM-2018-U-017433/DRM-2018-U-017433-Final.pdf?contentTicket=19gr23r22u4f1j1b6q4pv&Reload=1629242594111&__dmfClientId=1629242585249; Megan K. McBride, Zack Gold, and Kasey Stricklin, *Social Media Bots: Implications for Special Operations Forces*, CNA, 2020, accessed Aug. 17, 2021, https://www.cna.org/CNA_files/PDF/DRM-2020-U-028199-Final.pdf.

nontechnical language, the psychological mechanisms that disinformation campaigns exploit: initial information processing (IIP); cognitive dissonance (CD); group membership, belief, and novelty effects (GBN); and emotion and arousal effects (EA).[10]

Report organization

This report builds on the work contained in the primer; however, it transitions from theory to practice by focusing on real-world examples. After it presents an introduction, a brief discussion of disinformation, and a brief explanation of each mechanism, it focuses on five case studies that show these mechanisms operating in the real world. In bridging the gap between theory and practice, it illustrates why specific examples of disinformation are both spread and absorbed at high rates and which mechanisms are active in successful examples of disinformation. The case studies are robust disinformation campaigns executed by US adversaries, and address a range of domestic and international issues affecting US interests both at home and abroad.

The report's conclusion discusses near- and mid-term implications for national security. It also recommends ways to counter the spread of disinformation psychologically (as opposed to technologically or politically). In doing so, it addresses a vulnerability that has been underexplored in national security conversations.

[10] Wolters et al., *Psychology of (Dis)information*.

Disinformation: What It Is and Why We Care

Defining disinformation

As national and global concerns about disinformation campaigns become more mainstream—a shift precipitated in part by the 2016 US presidential election, and in part by the rise of COVID-19 disinformation—the definitions of the relevant terms may seem to be multiplying.[11] Particularly confusing is the interchangeability with which the terms *disinformation* and *misinformation* tend to be used. Experts recognize that these terms are easily differentiated by attention to the *intent of the content's creator* (Figure 1).

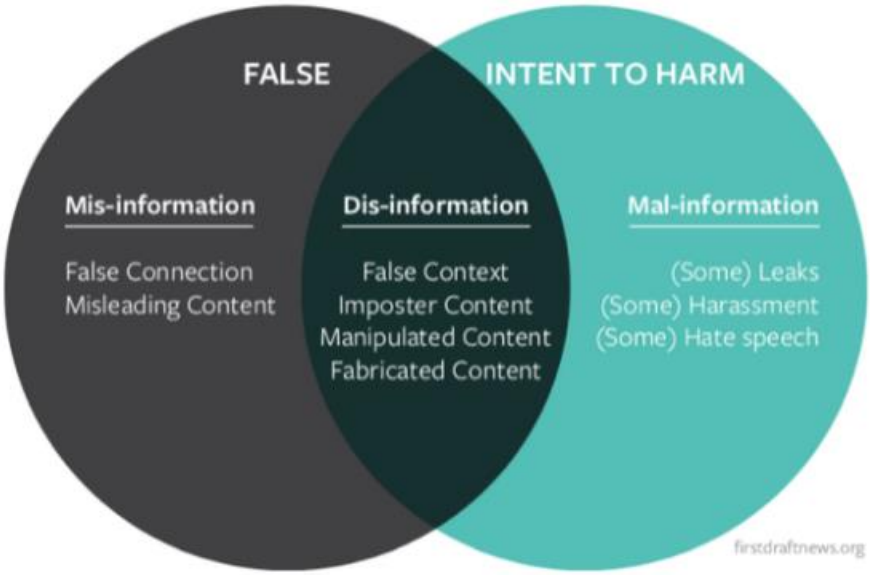
- *Disinformation* is information that is known to be false, and that is spread with the explicit goal of deceiving. In this case, the creator of the information (i.e., the tweet, the post, the email) *intends to deceive.*
- *Misinformation* is information that is false, but that is spread without a desire to deceive. In this case, the creator of the information (i.e., the tweet, the post, the email) *does not intend to deceive*.

This may seem like a relatively straightforward distinction, but in the real world this would require *knowing the intent* of the person who *created the content* you are viewing. And this is a nearly impossible task for the average user.

Because the average user does not know the intent of the content creator, this distinction is irrelevant for the purposes of our report. The object of analysis in this report is the *psychology of how disinformation works* (and not on, for example, adversary intentions), and so the distinction between disinformation and misinformation is not important. We usually don't know who created the content we consumed or what they intended to accomplish with it. As a result, from a psychological perspective, we engage with disinformation and misinformation in very similar ways.

[11] This section is a condensed version of a longer discussion available in our companion report.

Figure 1. Taxonomy of disinformation, misinformation, and mal-information



Source: Temir Asanov, "Fake News in Modern News Media: Disinformation, Misinformation and Malinformation," Mar. 17, 2019, https://medium.com/@tasanoff/fake-news-in-modern-news-media-disinformation-misinformation-and-malinformation-e4fdfa2ab571.

Why disinformation matters

The effect of disinformation

There is little to debate about the prevalence of disinformation in the modern world, particularly after the avalanche of COVID-19 disinformation that has been produced over the last 18 months. In fact, a 2018 MIT study of Twitter data found that disinformation was more successful than truth on social media by almost every known metric: it spread "significantly farther, faster, deeper, and more broadly than the truth in all categories of information."[12] What remains unclear is whether disinformation can actually change people's minds. Measuring an effect of this sort is extremely difficult because it is nearly impossible to know what someone might have believed or felt if they *hadn't* seen a piece of disinformation.

It is also very difficult to prove definitively that disinformation caused a change in behavior. In some instances, disinformation has very likely resulted in actions that have clear national

[12] Vosoughi, Roy, and Aral, "Spread of true and false news online."

security implications. The 2005–2006 rallies protesting the publication of cartoon depictions of Muhammad are an excellent example of disinformation resulting in political violence (including 100 deaths and 800 injuries). Though this case is rarely mentioned in conversations about disinformation (perhaps because it did not occur online), a significant driver for the rallies and violence was an informal publication known as the "Akkari-Laban dossier." This document contained inflammatory, false, and out-of-context information designed to anger those who read it. As just one example, it contained a picture of a man dressed as a pig and implied that the image was a Western depiction of Mohammad (the image was an Associated Press photo of a French pig-squealing contest).[13] The dossier did contain accurate information, but the admixture of accurate and inaccurate data makes it a clear example of disinformation. It is impossible to know whether the dossier was the direct cause of the rallies, violence, injuries, and death; however, it is widely recognized that the dossier exacerbated an already delicate situation and thus contributed meaningfully to the tragedy that followed.

Disinformation and our adversaries

It is possible that research will eventually show that today's disinformation is not changing the beliefs, attitudes, or behaviors of those who consume it. Even if it does, however, we should not assume that this will always be the case, because the production of disinformation will continue to evolve. This is especially concerning, given that disinformation is being produced not merely by individuals with personal agendas but also by state and nonstate US adversaries.

The most commonly cited goals for adversary use of disinformation include causing chaos and confusion, sowing discord, distracting from an issue, casting doubt, and making the truth seem unknowable.[14] However, even though these short-term objectives seem threatening in themselves, they may also be part of a broader long-term strategy for gaining global influence, diminishing US influence, claiming great power status, securing a regime, and more.[15] As states have become aware of the benefits of using information for achieving these objectives (including the ease with which they can harness the power of social media, the inexpensive nature of these efforts, and the relatively low risk of engaging in such behavior), information influence has come to hold a more prominent place in their doctrine, strategy, and thinking.

A primary US adversary in this space, Russia, has come to view the information space as one of the foundational areas in which states compete today.[16] Russia views information

[13] Martin Asser, "What the Muhammad Cartoons Portray," BBC News, Jan. 2, 2010, http://news.bbc.co.uk/2/hi/middle_east/4693292.stm.

[14] Dean Jackson, "Issue Brief: How Disinformation Impacts Politics and Publics," National Endowment for Democracy, May 29, 2018, https://www.ned.org/issue-brief-how-disinformation-impacts-politics-and-publics/.

[15] Kasey Stricklin, "Why Does Russia Use Disinformation?," *Lawfare,* Mar. 29, 2020, https://www.lawfareblog.com/why-does-russia-use-disinformation.

[16] Kliman et al., *Dangerous Synergies*.

confrontation as integral to traditional military capabilities, with utility during every phase of conflict, including during peacetime. In fact, it is thought that the use of such nonmilitary means as information can preclude an armed conflict by allowing Moscow to shape decision-making and internal dynamics within an adversary state, creating conditions favorable to Russia.[17]

Another US adversary in this space, China, has long sought to exercise control over information that circulates, and has recently begun to extend its influence beyond its borders.[18] The Chinese government's efforts were typically overt propaganda efforts to push Beijing's preferred narrative,[19] but after the onset of the coronavirus pandemic, China began to act more covertly by spreading disinformation on social media that obscured the virus's roots and shifted blame for the pandemic to the US.[20]

Iran's military also underscores the importance of informational control for offensive and defensive purposes alike;[21] a recent report described Iran's disinformation efforts as "public diplomacy under duress."[22] Iran sees itself as constantly defending itself against the information efforts of other countries,[23] and endeavors to set up a narrative structure that its adversaries cannot easily meddle with or take down. This includes positioning itself as a leader in the Muslim world and serving as a bastion against perceived US and Western regional intervention.[24]

The US government views disinformation in a patently different way, and is grappling with a range of ethical and legal implications that are not issues for its adversaries in this space. As a result, it still does not have a clear strategy for countering information operations, and its efforts in this space tend to be largely reactive.[25] The study of disinformation is important to ensure both that the US understands the role of disinformation in adversaries' strategies and that the US does not find itself unwittingly behind or vulnerable in a key area of competition.

---

[17] Kliman et al., *Dangerous Synergies*.

[18] Kliman et al., *Dangerous Synergies*, p. 5.

[19] Sarah Cook, "Welcome to the New Era of Chinese Government Disinformation," *The Diplomat,* May 11, 2020, https://thediplomat.com/2020/05/welcome-to-the-new-era-of-chinese-government-disinformation/.

[20] Joshua Kurlantzick, "How China Ramped Up Disinformation Efforts During the Pandemic," Council on Foreign Relations, Sept. 10, 2020, https://www.cfr.org/in-brief/how-china-ramped-disinformation-efforts-during-pandemic.

[21] "IRGC (Islamic Revolutionary Guard Corps)," Counter Extremism Project, accessed Feb. 3, 2021, https://www.counterextremism.com/threat/irgc-islamic-revolutionary-guard-corps; Emerson T. Brooking and Suzanne Kianpour, *Iranian Digital Influence Efforts*, Atlantic Council, 2020, https://www.atlanticcouncil.org/wp-content/uploads/2020/02/IRAN-DIGITAL.pdf.

[22] Brooking and Kianpour, *Iranian Digital Influence Efforts*.

[23] Brooking and Kianpour, *Iranian Digital Influence Efforts*.

[24] Brooking and Kianpour, *Iranian Digital Influence Efforts*.

[25] Doowan Lee, "The United States Isn't Doomed to Lose the Information Wars," *Foreign Policy,* Oct. 16, 2020, https://foreignpolicy.com/2020/10/16/us-election-interference-disinformation-china-russia-information-warfare/.

Finally, it would be naïve to write a paper on disinformation—especially one grappling with its real-world effects—without acknowledging that domestic US actors are active in this space as well. Their objectives are also varied; prominent examples are attempts to shift political discourse on controversial issues (such as election integrity, medical advice, and climate change) or to garner support (measured in the form of votes, subscriptions, or clicks). The psychological mechanisms that allow disinformation to spread are the same, whether an actor is international or domestic, benevolent or malicious.

The Psychology of Disinformation

Humans are constantly processing information that comes in through our senses.[26] Encoding, cataloging, deciding on action, and storing all that information could be a full-time job without mechanisms to process it all efficiently. Indeed, we don't have the "computing power" to appropriately process each new piece of information as entirely novel. In order to process data without overwhelming our capacity to do so, our brains rely on a series of adaptive mechanisms to triage and organize information as quickly and efficiently as possible.

This report explores the ways in which these mechanisms can be exploited to facilitate the spread of disinformation. Specifically, it identifies four key psychological mechanisms that are central to this dynamic. In doing so, it builds on a companion report—*The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*—that involved an extensive literature review across multiple domains, including disinformation, psychology, military science, foreign affairs, economics, computer science, and marketing.[27] Here we briefly define each mechanism and provide in-depth and real-world examples to illustrate the principles. A more comprehensive explanation can be found in the companion report.[28]

The four key psychological mechanisms related to the absorption and spread of disinformation are: initial information processing (IIP); cognitive dissonance (CD); the influence of groups, beliefs, and novelty (GBN); and the role of emotions and arousal (EA). Table 1 on the following page summarizes these psychological principles, provides a brief explanation, and shows the connection between that concept and disinformation.

[26] This section is a condensed version of a longer discussion available in our companion report.

[27] Wolters et al., *Psychology of (Dis)information*.

[28] Wolters et al., *Psychology of (Dis)information*.

Table 1. Psychological mechanisms relevant for disinformation adoption and spread

Mechanism	Explanation	Application to disinformation
Initial information processing	We process information as efficiently as possible and that can make us vulnerable to mistakes.	We can accept disinformation as true because we aren't thinking deeply and critically.
Cognitive dissonance	When we are confronted with something that goes against our beliefs, we are motivated to resolve the conflict.	We accept disinformation that supports our initial beliefs and try to reject information that disconfirms our initial beliefs.
Group, belief, and novelty	We more readily share information with people with whom we identify, when we believe it is true, and when it is novel or urgent.	We accept and share disinformation more readily when it comes from people we know, when it appeals to what "our group" believes, and when we think it is new.
Emotions and arousal	We pay more attention to information that makes us feel positively or arouses us to act.	We are more likely to share disinformation if it is constructed to elicit high-arousal emotions.

Source: CNA.

A critical takeaway from the identification of these mechanisms is that they are not unique to absorbing and spreading disinformation. These same principles are key to absorbing and spreading true information as well. Thus, at an individual level, it appears that disinformation is absorbed and spread through normal, routine, and adaptive mechanisms, which malign actors can exploit and manipulate for their own objectives.

Initial information processing (IIP)

Our mental "processing capacity" is limited; we simply cannot deeply attend to all new information we encounter. Our brains take mental shortcuts to incorporating new information, and those shortcuts can open us up to mistakes. To the extent that we do not process information as deeply as we should, disinformation can be construed as true information. We process information through two mechanisms: automatic and controlled processes. Automatic processes are fast, are effortless, and use very few cognitive resources. If something is familiar enough, we are unlikely to think deeply about it. For example, if someone drives the same road to work every day and comes upon their exit sign, they automatically turn on their signal and proceed to the exit ramp. This action requires very little thought. Using very few cognitive resources to process this information and act allows us to use reserved cognitive resources for other things. Controlled processes are slower, are more deliberate, and require considerably more resources and energy. Going back to the driving example, if the exit you were expecting to see was closed, your brain would begin controlled processes to determine alternative routes

for getting to work. Although the theory posits two processes, researchers acknowledge that there is really a gradient from the most automatic to the most controlled processes.[29]

Cognitive dissonance (CD)

Cognitive dissonance describes the discomfort we feel when we are confronted with two competing ideas. We are motivated to reduce the dissonance by changing one attitude, removing (ignoring) the contradictory information, discounting the importance of contradictory information, or increasing the importance of compatible information.[30] Festinger's classic example is of smokers encountering information indicating that smoking is bad for their health. In this case, the smoker has four options:

1. Change behavior or adopt new attitude (e.g., stop smoking).
2. Continue to believe that smoking is not bad for health (remove the contradictory information).
3. Compare risk from smoking to risk from something worse, such as auto accidents (reducing the importance of opposing information).
4. Think about the enjoyment of smoking and its good effects, such as nicotine's ability to suppress appetite (increase the importance of compatible information).

If disinformation supports our initial beliefs or creates less dissonance than true information, we are more likely to believe the disinformation.

Influence of group membership, beliefs, and novelty (GBN)

Not all information is equally valuable to individuals. Our group memberships, our beliefs, and the uniqueness of information each influence whether we absorb information and share it with people in our social networks. We are more likely to share content with people we consider members of our group (G), when we believe the information is true (B), and when we assess information to be novel or urgent (N). The Group, Belief, Novelty (GBN) model helps explain the likelihood that someone will pass content (rumors specifically) on to others.[31] This model explains, to some degree, the circulation of some urban myths online. In a case such as this, someone might receive an email from an acquaintance who has minimal internet savvy but is a member of our group (G); the email's content will likely be written in a way that suggests truth (B); and the email will contain information that is important or critical to someone's

[29] S.T. Fiske and S.E. Taylor, *Social Cognition: From Brains to Culture* (United Kingdom: SAGE Publications, 2016).

[30] L. Festinger, *A Theory of Cognitive Dissonance* (Evanston, Illinois: Row, Peterson, 1957).

[31] B.P. Brooks, N. DiFonzo, and D.S. Ross, "The GBN-dialogue model of outgroup-negative rumor transmission: Group Membership, belief, and novelty," *Nonlinear Dynamics, Psychology, and Life Sciences* 17, no. 2 (2013).

safety (N). One example, included below, is the urban legend of the "killer in the backseat," which has been circulating on the internet for years.

Figure 2. Internet Urban Legend

[Collected on the Internet, 1999]

Snopes

I am passing this along because I know of a incident similar to this. My girlfriend was getting some gas and when she attempted to return to her car the gas station attendant starting yelling at her and telling her she did not pay yet. When she went back in to argue about having already paid the attendant told her he just wanted to get her back in because he saw someone crawl in the backseat of her car, and that he had already called the police. So it's worth taking to heart.

This is a true story. it has been "ritual" of gang members to take one body part from women as an initiation into gangs. the rule is that it has to be in a well lit area and at a gas station, so be careful. they tend to lay under the car, and slash females' ankles when she goes to get in her car, causing her to fall and then they cut off a body part and roll and run. they are known to hide behind the gas pumps too, so be careful. It might sound bizarre and gross, but the bigger the body part the higher the initiation they receive.

Source: Barbara Mikkelsen, "The Killer in the Backseat," Urban Legends Reference Pages, http://www.snopes.com/horrors/madmen/backseat.htm.

Emotion and arousal (EA)

Just as not all information is equally valuable, not all information affects us the same way. Research demonstrates that we pay more attention to information that makes us feel good or that arouses us to act. In fact, marketing campaigns rely on the effect a message has on a person's emotions and subsequent behavior to buy products, donate to a cause, or vote for a candidate. Countless books, courses, and practices have been built on the concept of using messages to influence people's perceptions and behaviors, including the ubiquitous *How to Win Friends and Influence People* by Dale Carnegie and the extensively researched *Influence:*

Science and Practice by Robert Cialdini.[32] Although a message has multiple compelling aspects that can affect its ability to persuade, people are more likely to respond to (and share) information that is interesting, elicits positive emotions, or arouses action.[33] That means we are more likely to share information if we feel awe, amusement, or anxiety than if we feel sadness or contentment. In short, disinformation is more likely to be absorbed and shared if it is constructed to be emotional and arousing.

[32] D. Carnegie, *How to win friends and influence people* (Simon & Schuster, 1936); R. Cialdini, *Influence: Science and Practice*, 4th ed. (Boston: Allyn and Bacon, 2001).

[33] Jonah Berger, "Arousal Increases Social Transmission of Information," *Psychological Science* 22, no. 6 (2011), https://journals.sagepub.com/doi/10.1177/0956797611413294.

Disinformation in the Real World

Because disinformation is produced and spread every day, there are near endless examples for analysis. We selected the case studies below because they are examples of disinformation campaigns executed by US adversaries; in aggregate, they address a wide variety of domestic and international issues affecting US interests both at home and abroad; and they illustrate the full range of psychological mechanisms that this report addresses.

In each of the cases below we explain what happened—providing multiple images from the disinformation campaign itself—and then identify the specific psychological mechanisms that may have contributed to the campaign's longevity or promulgation, as summarized in Table 2.

Table 2. Psychological mechanisms analyzed in each case study

	IIP	CD	GBN	EA
Operation Secondary Infektion	✓		✓	✓
Endless Mayfly	✓			
US Presidential Election		✓		✓
Hong Kong: Protestors or Terrorists?		✓		✓
#CoronaJihad			✓	✓

Source: CNA.

Many times, pieces of all four components are at play, so it is difficult to state definitively that fewer mechanisms are present. We have focused our analysis on what we assess to be the primary mechanisms operating in each example.

Operation Secondary Infektion

Primary mechanisms: IIP, GBN, and EA

Background

Over the course of six years, beginning in 2014, an unknown Russian entity mounted a complex, large-scale disinformation campaign directed primarily at audiences in Western Europe and North America. What became known as "Operation Secondary Infektion" used social-media giants such as Google, Facebook, Reddit, Twitter, and YouTube to plant fake news stories, forged documents, and divisive content. Even more important to the operation were smaller platforms such as Medium, the German platform homment.com, and the California-

based indybay.org, which have minimal or no transparency requirements and allow users to easily create false profiles and post content.[34]

Operating in nine languages and on more than 30 social networks and blogging platforms, Secondary Infektion sought both to foment discord and division within countries and institutions that Russia perceived as its adversaries, and to arouse suspicions about critics of the government of Vladimir Putin. Toward those ends, the campaign worked to sow and nurture a set of toxic narratives and counternarratives. As depicted by Secondary Infektion, Europe was divided and weak; the United States and its allies were interfering aggressively in the affairs of other nations; "invading" Muslims were threatening to overrun Central and Western Europe; and opponents of the Putin regime were dissolute, corrupt, or unhinged.[35]

According to the Atlantic Council's Digital Forensic Research Laboratory (DFRLab), Secondary Infektion's operators consistently employed a three-step process:

> They would create an account on an online platform and use it to post a false story, often incorporating forged documents. A second set of fake accounts would post expanded versions of the same story in multiple languages, using the original posts as their source. In the third step, additional fake social media accounts amplified the false stories and tried to bring them to the attention of the mainstream media.[36]

Given the scale and duration of Secondary Infektion, it is impossible to give a complete account of the operation in this case study. Instead, this case study will focus on components of the campaign in the United Kingdom, a major Russian target.

IIP/GBN: Assassination plots, forged documents, and the mainstream media

The disinformation operation targeting the UK consisted of astonishing stories involving UK political figures. One particular effort involved a fanciful narrative about an alleged plot to assassinate Boris Johnson, the recently resigned foreign minister (and future prime minister) who was a leading supporter of withdrawal from the European Union. An August 8, 2018, post on a Spanish-language Facebook account set up by the campaign's operators claimed that a

---

[34] Nika Aleksejeva et al., *Operation "Secondary Infektion": A Suspected Russian Intelligence Operation Targeting Europe and the United States*, Digital Forensic Research Laboratory, 2019, accessed June 8, 2021, https://www.atlanticcouncil.org/in-depth-research-reports/report/operation-secondary-infektion/; Ben Nimmo et al., *Exposing Secondary Infektion*, Graphika, 2020, accessed July 9, 2021, https://secondaryinfektion.org/; Secondary Infektion was named after "Operation Infektion," a Soviet-era disinformation campaign which claimed that the US military had created the AIDS virus. Adam Rawnsley, "Russian Trolls Hype Coronavirus and Giuliani Conspiracies," *Daily Beast*, Apr. 9, 2020, accessed July 13, 2021, https://www.thedailybeast.com/russian-trolls-hype-coronavirus-and-giuliani-conspiracies;

[35] Ben Nimmo et al., *Exposing Secondary Infektion*, Graphika, 2020, accessed July 9, 2021, https://secondaryinfektion.org/.

[36] Aleksejeva et al., *Operation "Secondary Infektion*."

Spanish intelligence service had uncovered a plot by "radical opponents of Brexit" to assassinate Johnson (Figure 3).[37] The post on the fake Facebook account included a letter, purportedly from the Spanish foreign secretary to a prominent Spanish parliamentarian. A close inspection revealed the letter to be an obvious forgery—among other things, the foreign secretary's name was misspelled—but this wouldn't necessarily be clear to a casual reader.

Figure 3. Facebook post claiming the existence of an assassination plot against Johnson

MINISTERIO DE ASUNTOS EXTERIORES, UNIÓN EUROPEA Y COOPERACIÓN

El Ministro

D. Rafael Merino López
Presidente de la Comisión de Interior
Congreso de los Diputados
C/Floridablanca, s/n
28071 Madrid

Madrid, 1 de agosto de 2018

Estimado Sr. Merino:

Acuso recibo de su carta fechada el 30 de julio de 2018, en la que nos traslada la información sobre el posible atentado contra Boris Johnson preparado por los radicales oponentes del Brexit que buscan impedir que sea nombrado primer ministro del país.

Déjeme decirle en primer lugar que les agradezco a usted y a todos los miembros de la Comisión de Interior del Congreso de los Diputados su colaboración y la información que nos proporcionaron de manera oportuna. Comparto plenamente con usted la necesidad de informar a las autoridades del Reino Unido sobre el posible ataque contra señor Johnson lo antes posible. Estamos en contacto constante con el Ministerio de Asuntos Exteriores y de la Commonwealth del Reino Unido y mantenemos una cooperación muy estrecha, incluso en materia de seguridad. La lucha contra los movimientos radicales y terrorismo es una de las prioridades de la política de España. Creo que la única manera de erradicar tales actitudes y mantener la paz y la estabilidad en Europa es con el intercambio de información y trabajo conjunto de todos los países europeos.

Le puedo asegurar que dicha información será enviada por canales diplomáticos al Gobierno de Theresa May tan pronto como sea posible. Esperamos que las autoridades y fuerzas del órden británicas puedan garantizar la seguridad de Boris Johnson y encontrar a los presuntos autores del atentado.

Reciba un cordial saludo,

Josep Borrel Fontelles

Source: Aleksejeva et al., *Operation Secondary Infektion,* p. 19.
[a] The Facebook comment attached to the image read: "Radical Brexit opponents are preparing an assassination attempt on Boris Johnson."

[37] Aleksejeva et al., *Operation "Secondary Infektion*."

The next day, a Spanish-language article based on the forgery was shared on three Spanish forums. Four days later, an account called "Matt Porter" posted an English-language translation on Medium—a relatively mainstream US-based website on which a mix of professional and independent individuals publish content.[38] In another account created that same day, "portmatt" posed the article to defendingtruth.com, homment.com, debatepolitics.com, and other forums, before the account disappeared. Finally, using an account created just two days later, "Illinoiss" posted a meme that included the "Matt Porter" article (Figure 4).

Figure 4. "Illinoiss" meme

Source: Aleksejeva et al., *Operation Secondary Infektion*.

The assassination narrative attempted to exploit two psychological mechanisms. First, the nature of the threat—that is, the assassination of a high-ranking British government official—ensured that it would be received as urgent information (thus triggering an increased inclination to share based on the GBN mechanism). Second, the disinformation was designed

[38]Aleksejeva et al., *Operation "Secondary Infektion*."

to avoid careful scrutiny by presenting the data in formats and on forums that readers were likely to assess as trustworthy. In other words, the content exploited the tendency towards automatic (low-scrutiny) initial information processing mechanisms through the use of official-looking letterheads, signed documents, and placement on websites such as Medium. Close scrutiny might have revealed that the Spanish foreign secretary's name was misspelled, but the use of a formal letterhead and an official signature was a psychological hack carefully designed to ensure that we wouldn't look closely enough to notice any errors contained within.

GBN/EA: Attempted assassinations and terrorist groups

In another example, between March 2018 and April 2019 the Secondary Infektion operators attempted to sow discord and confusion within the UK by linking two highly inflammatory topics: the activities of the Real Irish Republican Army (Real IRA), a leading dissident Irish republican group, and the attempted assassination of a former Russian intelligence officer, Sergei Skripal, in Salisbury, England, on March 4, 2018. Just two weeks after the assassination attempt, some 15 news groups received an article posted to Medium by a Russian-linked Facebook account.[39]

The article made the startling allegation that Gavin Williamson, the UK defense minister at the time, had tweeted that the Real IRA had assisted in the attempt to kill Skripal. The article included a screenshot of the forged tweet, and claimed that "militants of the Real Irish Republican Army had transferred a component of the 'Novichok' nerve agent to Salisbury with the view to poison the former Russian military intelligence colonel" (see Figure 5). The phony account, moreover, embedded a genuine Sky News report along with the fake tweet.[40]

[39] Aleksejeva et al., *Operation "Secondary Infektion*." For more on dissident Irish Republican groups, see Martyn Frampton, *Legion of the Rearguard: Dissident Irish Republicanism* (Dublin and Portland, Oregon: Irish Academic Press, 2011).

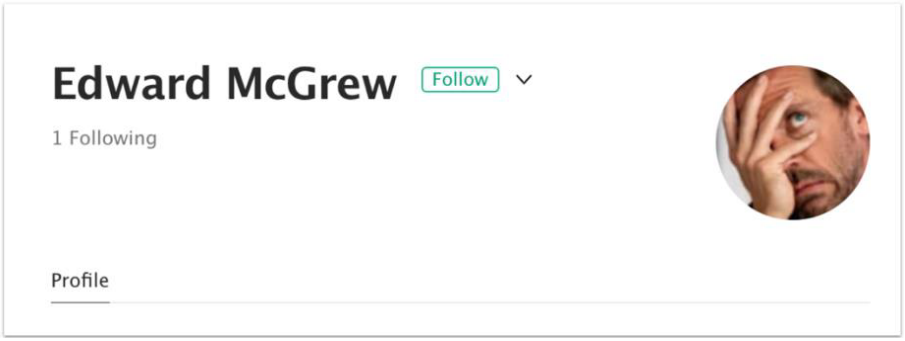
[40] Martin Robinson, "Russia Has Tried to Reignite the Troubles With Fake Social Media Posts," *Daily Mail*, June 26, 2019, accessed July 14, 2021, https://www.dailymail.co.uk/news/article-7182739/Russia-tried-reignite-Troubles-fake-social-media-posts.html.

Figure 5. Medium article planted by a Russian-linked Facebook account

Source: Medium: "Edward McGrew."

The creation and dissemination of the forged tweet had all the hallmarks of Operation Secondary Infektion. For example, the source account on Medium, named "Edward McGrew," posted just this single article and used a stolen profile picture of Hugh Laurie, the British actor and comedian (see Figure 6).

Figure 6. The "Edward McGrew" source account, with Hugh Laurie's picture, on Medium



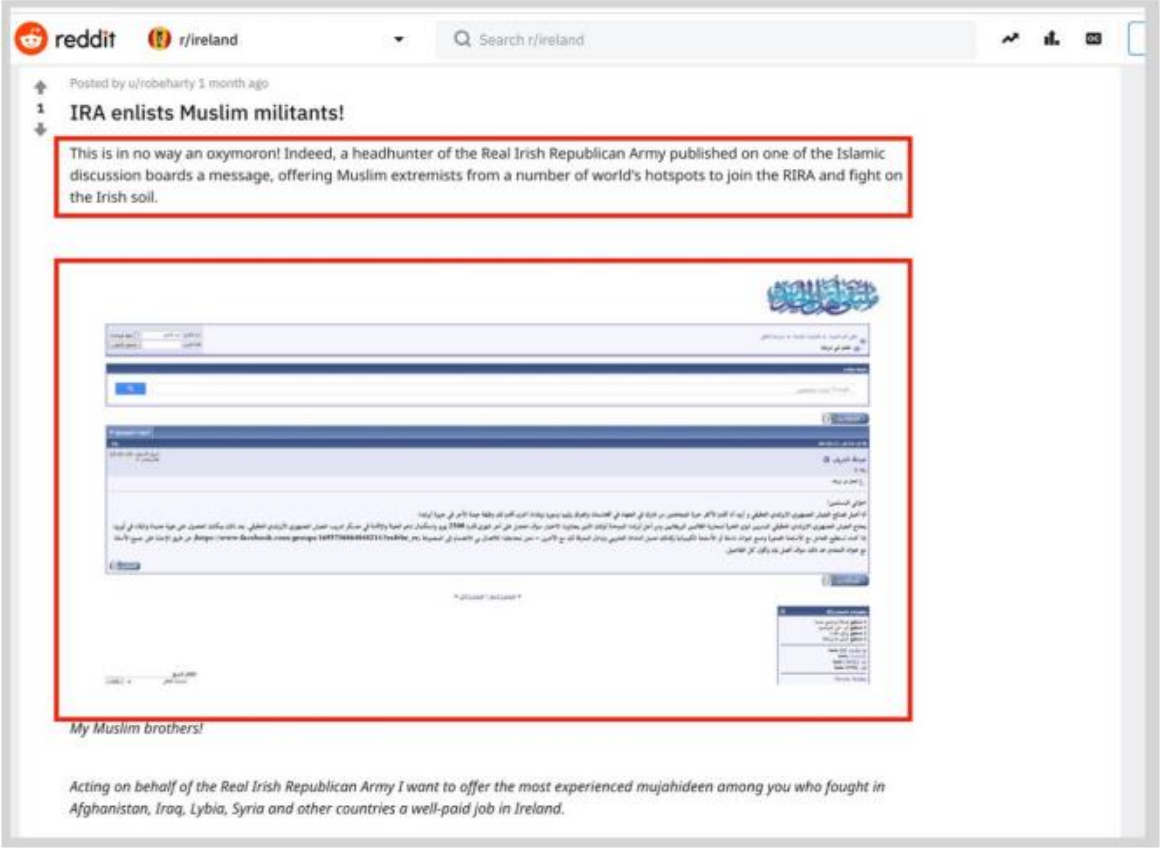
Source: Aleksejeva et al., *Operation Secondary Infektion*.

A final attempt effort to take advantage of the already inflammatory issue of the Real IRA occurred on April 22, 2019. In this case, an individual on the r/Ireland subreddit—a public forum for those interested in posting or discussing content about Ireland—asserted that the Real IRA was recruiting jihadists to wage war in Ireland (see Figure 7).[41] Using garbled syntax, the Reddit user, an individual claiming to be in Ireland but with no history of pre-existing posts, claimed that "a headhunter of the Real Irish Republican Army published on one of the Islamic discussion boards a message" encouraging Islamist fighters to join the Real IRA's fight in Ireland.[42]

[41] Aleksejeva et al., *Operation "Secondary Infektion*."

[42] *IRA enlists Muslim militants!*, Reddit, Reddit Post, Apr. 22, 2019.

Figure 7. r/Ireland subreddit post



Source: u/robeharty, "IRA enlists Muslim militants!," Reddit, 2019.

The Real IRA narrative differed significantly from the assassination narrative, but still attempted to exploit two psychological mechanisms. First, the nature of the threat—that is, content that involved attempted assassinations, violent dissident groups, and known terrorist groups—ensured that it would be received as urgent information (again triggering an increased inclination to share based on the GBN mechanism). Second, the topics were also both emotionally arousing, given the UK's long history of fighting with the IRA (and its modern incarnation, the Real IRA) and its more recent experience with violent jihadist groups such as Al Qaeda and ISIS. The information was, perhaps, less expertly packaged in that it wasn't presented in the form of a forged document or a leaked memorandum. This wasn't necessary, though, as it exploited other mechanisms to ensure its appeal and spread.

Endless Mayfly

Primary mechanisms: IIP

Background

The Iran-linked disinformation campaign "Endless Mayfly" is a network of false personas, social media accounts, and websites that impersonate legitimate news organizations. It has been active since at least 2016. The effort aligns with what is known of Tehran's approach to information operations, which the state has described as "content promotion" and has characterized as its most potent weapon against foreign adversaries.[43] Iran has also claimed that it has created numerous "cyber battalions" made up of more than 8,000 specialists to create and spread this content.[44] Iran's foreign disinformation operations are directed primarily against nations that it perceives to be its greatest enemies: Saudi Arabia, Israel, and the United States.

Endless Mayfly—a name given by researchers at the University of Toronto's Citizen Lab, which studied the network's operations from April 2016 through November 2018—unfolded in five overlapping phases, each of which involved the creation, dissemination, and amplification of false content.[45] Key narratives propagated during this effort included the following:

- Saudi Arabia's relations with the United Arab Emirates, France, the United Kingdom, and the United States were strained.
- Saudi Arabia's footprint in the Middle East was growing larger, thanks to cooperation between the kingdom and Muslim-majority countries.
- Saudi Arabia is a major supporter of terrorism.[46]

IIP: Fake websites and borrowed authenticity

The Endless Mayfly operation was ultimately reliant on a perception of authenticity that exploited the tendency towards low-scrutiny thinking during initial information processing. However, feigning authenticity required a few steps. The first step was the creation of personas, including fake students, activists, and reporters. Initially these persona accounts

[43] Emerson T. Brooking and Suzanne Kianpour, *Iran Digital Influence Efforts: Guerrilla Broadcasting for the Twenty-First Century*, Atlantic Council, 2020, accessed July 6, 2021, https://www.atlanticcouncil.org/in-depth-research-reports/report/iranian-digital-influence-efforts-guerrilla-broadcasting-for-the-twenty-first-century/.

[44] Brooking and Kianpour, *Iran Digital Influence Efforts*.

[45] Ronald Deibert, "Endless Mayfly: An Invasive Species in the Social Media Ecosystem," The Citizen Lab, May 14, 2019, accessed July 10, 2021, https://deibert.citizenlab.ca/2019/05/endless-mayfly/. The lab's director described Endless Mayfly as "an invasive species in the social media ecosystem."

[46] Gabrielle Lim et al., *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign*, 2019, accessed July 9, 2021, https://citizenlab.ca/2019/05/burned-after-reading-endless-mayflys-ephemeral-disinformation-campaign.

were relatively lean, consisting of little more than Twitter bios, but over time they grew more robust as the "users" acquired bylines on third-party websites such as *Buzzfeed* (see Figure 8).

Figure 8. Endless Mayfly persona's article published on Buzzfeed Community

Source: Lim et al., *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign.*

In steps two and three, Endless Mayfly created 72 "lookalike" domains that mimicked well-known, highly credible news outlets, including *The Guardian*, *The Atlantic*, *The Independent*, and *Haaretz*, and populated them with malicious, scurrilous, and false "copycat" content (Figure 9 and Figure 10).[47,48] People viewing this content were not likely to process the images deeply enough to realize they were not the same as "official" websites.

[47] Lim et al., *Burned After Reading.*

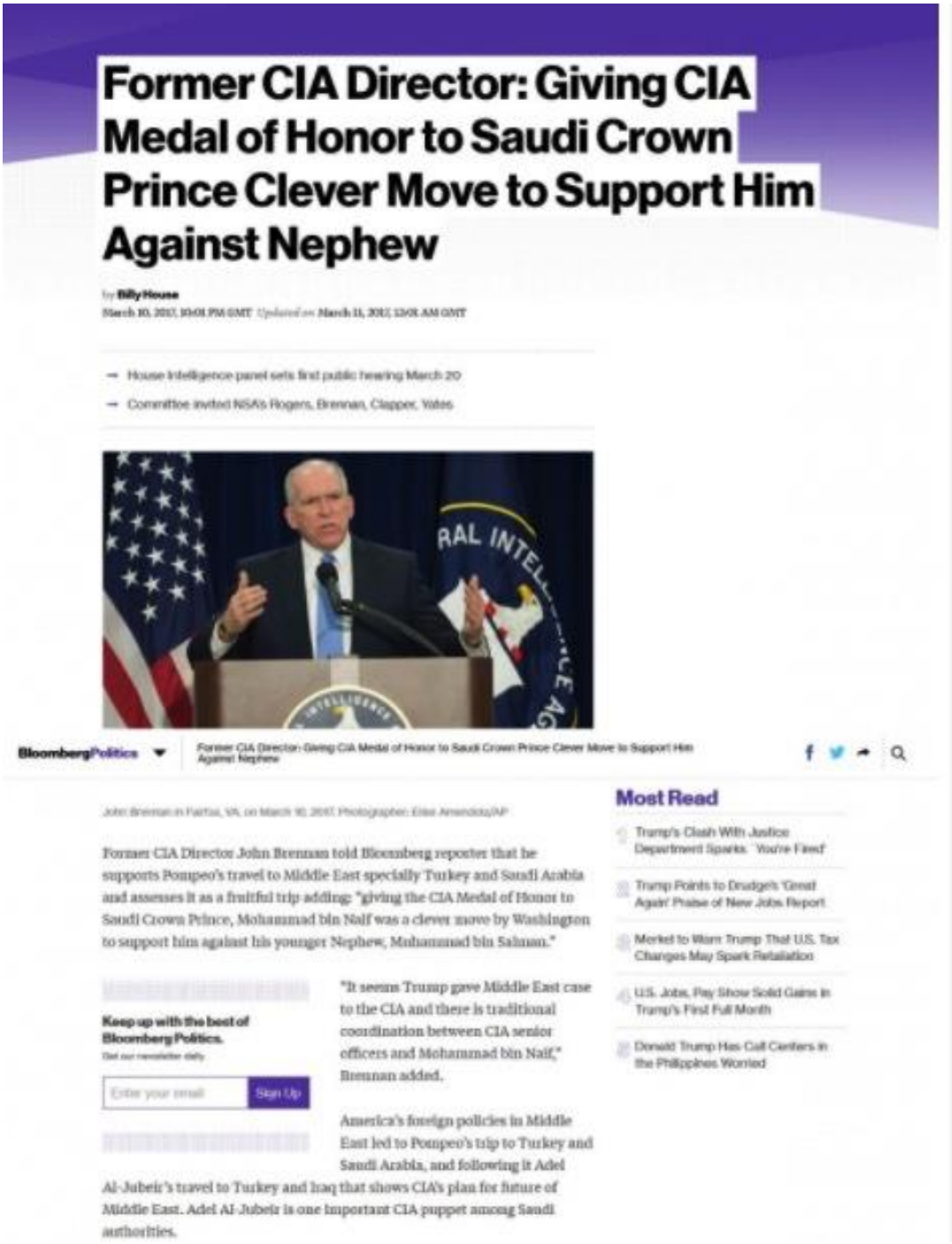
[48] Iran-linked operators have also employed a network of phony news sites (e.g., "Liberty Front Press" and "Instituto Manquehue") and associated accounts on multiple social media platforms to disseminate and amplify

Figure 9. *The Guardian* lookalike domain

Source: Screenshot of lookalike *Guardian* page via Craig Silverman and Jane Lytvynekno, "How A Hoax Made To Look Like A Guardian Article Made Its Way To Russian Media," Buzzfeed News, Aug. 15, 2017, accessed Aug. 6, 2021, https://www.buzzfeednews.com/article/craigsilverman/how-a-hoax-made-to-look-like-a-guardian-article-made-its.

pro-Teheran narratives. FireEye, "Suspected Iranian Influence Operations: Leveraging Inauthentic News Sites and Social Media Aimed at U.S., U.K, Other Audiences," FireEye, Aug. 21, 2018, accessed July 20, 2021, https://www.fireeye.com/blog/threat-research/2018/08/suspected-iranian-influence-operation.html.

Figure 10. *Bloomberg Politics* lookalike domain



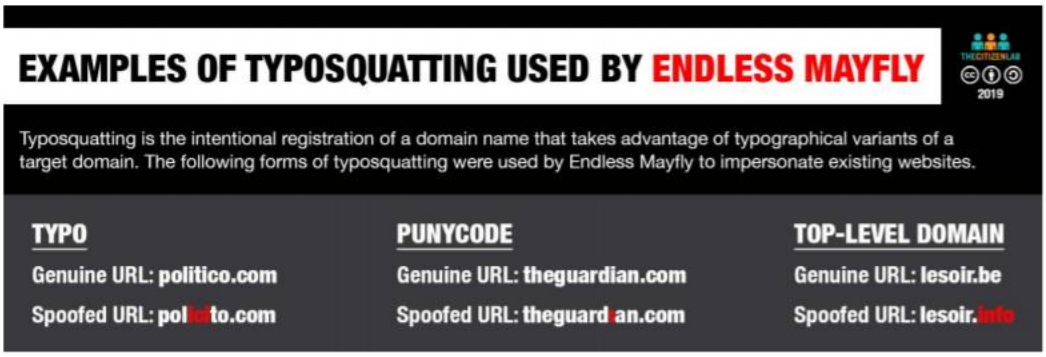
Source: Lim et al., *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign.*

To do this, Endless Mayfly operators used "typosquatting" to add further depth to the deception. According to the Center for Internet Security, typosquatting

> attempts to take advantage of typographical errors (i.e. "typos") introduced by users when URLs are typed directly into the address bar. By capitalizing on user error, cyber threat actors funnel unsuspecting users to illegitimate domains that closely mimic originals. This tactic involves the purchase and registration of domains similar to an existing domain.[49]

In the simplest example of this technique, a malicious actor might buy the domain www.nytime.com to take advantage of users who type too quickly (and leave out the 's'). Those who make this mistake might then be directed to a fake version of the *New York Times* that *appears to be real* but that includes articles with inaccurate and falsified content. In other instances, typosquatting takes advantage not of user errors (i.e., typos), but of user inattention by inserting non-standard characters into a URL and hoping that someone might not notice, or by purchasing a website on an alternative domain. The Citizen Lab illustrates all three examples in its report on Endless Mayfly (Figure 11).[50]

Figure 11. Types of typosquatting used in Endless Mayfly



Source: Lim et al., *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign*, p. 16.

In step four, the disinformation campaign worked to amplify the content it created. Endless Mayfly personas, posing as journalists, students, and activists, actively promoted content on Twitter, by posting screenshots of phony articles and sharing links to false articles with reporters and activists, many of whom also received private direct messages from the

[49] Center for Internet Security, "MS-ISAC Security Primer: Typosquatting," Center for Internet Security, Feb. 2018, accessed July 14, 2021, https://www.cisecurity.org/wp-content/uploads/2018/02/MS-ISAC-Typosquatting-Security-Primer.pdf.

[50] Lim et al., *Burned After Reading*.

disinformation operators (Figure 12).[51] In the judgment of Citizen Lab analysts, embedding screenshots of phony articles in the tweets helped increase dissemination: "Recent studies have shown that users of social media tend only to read headlines, and that they often share a link without reading the body of the article."[52]

Figure 12. Tweets including screenshots of a fake news story

Source: Lim et al., *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign*.

In Endless Mayfly's fifth and final stage, following successful amplification efforts on social media, operators removed the phony articles and redirected the links to the impersonated media sites it had created (i.e., visitors to the misspelled www.theguardıan.com would now be directed to the real website, www.theguardian.com). However, according to the Citizen Lab, references to the spurious articles remained online, "further creating the appearance of a legitimate story, while obscuring its origins."[53]

[51] Lim et al., *Burned After Reading*.

[52] Lim et al., *Burned After Reading*. For more on this point see Maksym Gabielkov, Arthi Ramachandran, Augustin Chaintreau, and Arnaud Legout, "Social Clicks: What and Who Gets Read on Twitter?" ACM SIGMETRICS / IFIP Performance 2016, June 2016, accessed July 14, 2021, https://hal.inria.fr/hal-01281190/document.

[53] Lim et al., *Burned After Reading.*

The overall effect of the Endless Mayfly campaign is difficult to assess. However, in the view of Citizen Lab, Endless Mayfly content did lead to uncertainty among journalists, and incorrect reporting. Moreover, even when stories were later proven to be false, "confusion remained about the intentions and origins behind the stories."[54] To a large degree, this occurred because Endless Mayfly cleverly exploited our tendency to trust information that comes from credible sources. In other words, Endless Mayfly borrowed the authenticity of legitimate news sources by making lookalike content, and in doing so effectively triggered our tendency to engage more superficially with data that are familiar (i.e., IIP). We might approach an unfamiliar news source with skepticism, asking ourselves whether it can be trusted, but we already know that *The Guardian* is a legitimate news source so this closer scrutiny isn't triggered (i.e., our automatic, not controlled, information-processing apparatus is activated). In masquerading as legitimate news sources, or planting stories on legitimate news sites, Endless Mayfly ensured that people would be predisposed to trust its content. The effort also, however, undermined trust in the very sites whose authenticity it borrowed, by leaving viewers and readers perplexed about what *really* happened.

The 2016 US presidential election

Primary mechanisms: CD, EA

Background

In July 2016, the FBI began to investigate an extensive Russian operation aimed at influencing the 2016 US presidential election by manipulating Americans and causing social divisions. These efforts turned out to be part of a three-pronged interference operation, and included an extensive, multi-year social media campaign that targeted a demographically diverse population from a wide range of interest groups and political ideologies with disinformation, memes, and divisive content.[55] A Russian state-supported group known as the Internet Research Agency (IRA) was identified as one of the primary actors behind this campaign.[56] Russian actors worked across all major social media platforms, including Facebook, Twitter, Instagram, YouTube, Reddit, Pinterest, Tumblr, Meetup, and Vine.[57] Perhaps as critical, though, was that Russian actors had a deep understanding of America's social media ecosystem and were able to mimic legitimate content—posting event invitations, memes, and news articles; linking across platforms; and sharing authentic and nonpolitical content. This not only made

[54] Lim et al., *Burned After Reading*.

[55] Renee DiResta et al., *The Tactics & Tropes of the Internet Research Agency*, New Knowledge, 2019, p. 10, accessed July 1, 2021, DigitalCommons@University of Nebraksa - Lincoln, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1003&context=senatedocs.

[56] DiResta et al., *Tactics & Tropes*, p. 99.

[57] DiResta et al., *Tactics & Tropes,* p. 5.

their pages more believable, but also ensured that posts appeared normal in the flow of users' social media feeds.[58]

EA: Flaming American anger

Much of the IRA's content focused on contentious issues such as race, patriotism, immigration, gun control, and LGBT rights, striving to stoke fear, anger, and division among Americans.[59] In other words, the content was explicitly designed to ensure emotional arousal. The IRA relied heavily on the use of images and memes—culturally relevant and broadly resonant images with accompanying text—to connect with target audiences.[60] This format ensured that the content would be easily digestible, and was likely chosen in hopes that it would increase shares (Figure 13).[61] Unsurprisingly, divisive posts on controversial issues resonated with audiences predisposed to receive them favorably, and the IRA's five most shared and liked posts focused on contentious and polarizing issues.[62]

[58] Young Mie Kim, *Uncover: Strategies and Tactics of Russian Interference in US Elections*, Project DATA, 2018, p. 3, accessed July 8, 2021, https://journalism.wisc.edu/wp-content/blogs.dir/41/files/2018/09/Uncover.Kim_.v.5.0905181.pdf.

[59] Young Mie Kim, *New Evidence Shows How Russia's Election Interference Has Gotten More Brazen*, 2020, accessed July 7, 2021, https://www.brennancenter.org/our-work/analysis-opinion/new-evidence-shows-how-russias-election-interference-has-gotten-more.

[60] Zakem, McBride, and Hammerberg, *Exploring the Utility of Memes*.

[61] Dan Keating, Kevin Schaul, and Leslie Shapiro, "The Facebook ads Russians targeted at different groups," *Washington Post*, Nov. 1, 2017, https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/.

[62] Philip N. Howard et al., *The IRA, Social Media and Political Polarization in the United States, 2012-2018*, Computational Propaganda Research Project, 2019, p. 8, accessed July 5, 2021, DigitalCommons@University of Nebraska - Lincoln, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1004&context=senatedocs.

Figure 13. Examples of emotionally arousing memes created and shared by the IRA

Source: DiResta et al., *The Tactics & Tropes of the Internet Research Agency,* p. 45.

The effect of the campaign on the election itself is unclear, but the IRA demonstrated a remarkable ability to produce content that was appealing enough to share. In fact, IRA Facebook posts alone were shared over 31 million times.[63] Unlike the previous case studies, this was not accomplished by creating content that pretended to be authentic or trustworthy. By contrast, the content in this campaign was not "fact" based but emotion based. As the research shows, content that is emotionally arousing—that calls us to action, increases anxiety, etc.—is more persuasive than information that is emotionally neutral. IRA effectively hacked this psychological mechanism by choosing topics that were already emotionally arousing. As Congressman Adam Schiff noted in November 2017, this campaign aspired to "sow discord by inflaming passions on a range of divisive issues."[64] Schiff went on to note that "Russia exploited real vulnerabilities that exist across online platforms," but it is perhaps equally accurate to say that Russia exploited real vulnerabilities in human psychology.[65]

CD: Affirming pre-existing beliefs

Further amplifying the effect of this incendiary content was the IRA's exploitation of Facebook's advertising features. The ability to microtarget specific segments of the population allowed the IRA to provide content, promote groups, and advertise events based on the interests, locations, views, and racial identities of the targeted sub-populations.[66] In other words, the IRA was able to provide *customized* content that affirmed their pre-existing beliefs (i.e., CD).

As one example, the IRA was able to target individuals who appeared to be against Muslim immigration based on their Facebook usage. These individuals received ads for groups such as *Stop All Invaders* and posts like the one below, uploaded in January 2017 (Figure 14), which asserted that President Obama had refused to ban Sharia Law and urged President Trump to take immediate action. This particular post was the IRA's most successful post featuring Trump on Facebook: it had the highest engagement, receiving 312,632 shares from presumably real users.[67]

[63] Howard et al., *IRA, Social Media and Political Polarization*.

[64] "Exposing Russia's Effort to Sow Discord Online: The Internet Research Agency and Advertisements," US House of Representatives Permanent Select Committee on Intelligence, accessed Sept. 27, 2021, https://intelligence.house.gov/social-media-content/.

[65] "Exposing Russia's Effort to Sow Discord Online: The Internet Research Agency and Advertisements."

[66] Keating, Schaul, and Shapiro, "Facebook ads Russians targeted at different groups."

[67] DiResta et al., *Tactics & Tropes*, p. 24.

Figure 14. *Stop All Invaders* post on Sharia Law

"On june 27, 1952, American government passed a law, called '1952 McCarran Walters act', that actually outlawed Sharia, but Obama never intended to enforce it or even let you know about it at all. Instead he started to import thousands of aggressive Muslim 'refugees', who refused to integrate and demanded to be allowed to live under sharia law instead of American constitution. Now it's within Donald Trump's authority to enforce that law and ban Sharia in every state across America. Do you want him to do that?" – Stop A.I., posted 1/23/2017

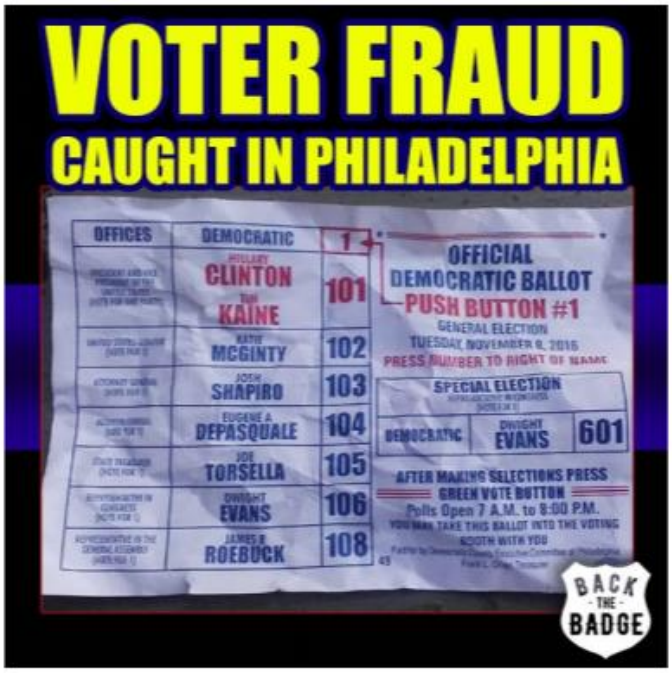
# Likes	# Comments	# Reactions	# Shares	Total Engagement
162850	93231	11014	312632	579727

Source: DiResta et al., *The Tactics & Tropes of the Internet Research Agency*, p. 24.

In another example, understanding the US right-wing audience's fears of voter fraud, the IRA focused at least 109 posts on the issue (Figure 15), amplifying claims about states interfering to help Hillary Clinton win or about "illegals" voting multiple times with aid from the Democratic Party.[68]

[68] DiResta et al., *Tactics & Tropes,* pp. 24, 77.

Figure 15. Meme claiming voter fraud



Source: DiResta et al., *The Tactics & Tropes of the Internet Research Agency*, p. 77.

And in a final example, the IRA focused on Black Americans in an attempt to exacerbate divisions within the community, and stoke distrust of Clinton and democratic institutions.[69] The group strategically exploited real tensions and preexisting grievances in an effort to divert Black voters' political energy away from the election and towards structural inequalities.[70] Posts focused on topics such as Black identity, police violence, the Black Lives Matter movement, incarceration, poverty, and white supremacy. The IRA-created Facebook page "blacktivist," for example, almost exclusively shared videos of police brutality against African Americans—some legitimate and others fabricated.[71] The IRA also attempted to take advantage of the increasingly popular Black Lives Matter movement; they created a site called "Black Matters US," and asserted that it was a more radical version of Black Lives Matter.[72]

[69] Kim, *New Evidence Shows How Russia's Election Interference Has Gotten More Brazen*.

[70] P. R. Lockhart, "How Russia exploited racial tensions in America during the 2016 elections," *Vox*, Dec. 17, 2018, accessed July 7, 2021, https://www.vox.com/identities/2018/12/17/18145075/russia-facebook-twitter-internet-research-agency-race.

[71] "Russian trolls' chief target was 'black US voters' in 2016," *BBC*, Oct. 9, 2019, accessed July 16, 2021, https://www.bbc.com/news/technology-49987657.

[72] Casey Michel, *Website targeting black Americans appears to be elaborate Russian propaganda effort*, ThinkProgress, 2017, accessed July 20, 2021, https://thinkprogress.org/black-matters-us-site-90625b18f262/.

Thirty IRA-supported Facebook groups targeting Black Americans amassed 1.2 million followers.[73]

Similarly, YouTube duo *William and Kalvin* appeared to be legitimate Black content creators from Atlanta who called Hillary Clinton a criminal and raised questions about the relationship between her and the Clinton Foundation (Figure 16). In reality, William and Kalvin were working under the funding, support, and direction of the Kremlin.[74]

[73] Scott Shane and Sheera Frenkel, "Russian 2016 Influence Operation Targeted African-Americans on Social Media," *New York Times*, Dec. 17, 2018, accessed July 11, 2021, https://www.nytimes.com/2018/12/17/us/politics/russia-2016-influence-campaign.html.

[74] Kim, *Uncover: Strategies and Tactics of Russian Interference in US Elections*, p. 11.

Figure 16. Post from William & Kalvin

Source: Kim, *Uncover: Strategies and Tactics of Russian Interference in US Elections*, p. 9.
The text reads: "Power to the People! We have to grow up, we have to wise up. We don't have any other choice this time but boycott the election. This time we choose between two racists. No one represents Black people. Don't go to vote. Only this way we can change the way of things…."

Though these topics are also emotionally arousing, they are not *merely* emotionally arousing. Instead, this content is designed to target specific sub-populations with messages that they are pre-disposed to accept because it affirms pre-existing beliefs (i.e., CD). As the research shows, we are more likely to believe information that supports pre-existing beliefs (e.g., that only President Trump will protect America from Islamic law, that voter fraud is a massive problem, that structural racism and police violence prevent Black Americans from achieving the

American dream). Thus, by developing targeted content and providing it to specific sub-populations, the IRA was able to inflame the anger of Americans at both ends of the political spectrum. This content didn't challenge the beliefs of those who received it; for the most part, far-right actors didn't receive IRA posts about police brutality, and far-left actors didn't receive IRA posts about illegal immigrants. Instead, this fabricated content was nearly indistinguishable from the organic content that filled their feeds. And, unfortunately, people were predisposed to receive it favorably because it affirms what they already held to be true (i.e., CD).

Hong Kong: protestors or terrorists?

Primary mechanisms: EA, CD

Background

In June 2019, the Chinese Communist Party (CCP) passed a new law allowing the extradition of Chinese citizens living in territories abroad, including Hong Kong. The extradition bill was seen by some as a Chinese attempt to repress Hong Kong's civil liberties. Protests began in Hong Kong, ultimately growing to include over one million people.[75] The protests started peacefully, but clashes with the police became more common over time.[76]

Between June and August 2019, in response to these protests, the Chinese government launched a coordinated disinformation campaign. The campaign started modestly as Chinese state-sponsored media initially posted a few articles condemning the violence and encouraging Hong Kong to return to the rule of law. However, following the vandalizing of the Legislative Council building and the occupation of the Hong Kong airport by protestors, the Chinese government began trying to more actively control the narrative around the demonstrations, leveraging propaganda and disinformation.[77]

The influence campaign was global, targeting Chinese citizens, members of the Chinese diaspora internationally, Hong Kong residents, and the international community.[78] It was one of China's first disinformation campaigns directed towards a broader international audience.[79] Disinformation accused the US and other Western powers of being behind the protests,

[75] Renee DiResta et al., *Telling China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives*, Stanford Internet Observatory, 2020, p. 21, accessed July 18, 2021, https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/sio-china_story_white_paper-final.pdf.

[76] Diresta et al., *Telling China's Story*.

[77] Diresta et al., *Telling China's Story*, p. 21.

[78] John Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," *China Brief* 19, no. 16 (2019): p. 3, accessed July 20, 2021, https://jamestown.org/wp-content/uploads/2019/09/Read-the-09-06-2019-CB-Issue-in-PDF.pdf.

[79] Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong,"

exaggerated the protesters' violence, and argued that protestors were trying to start an independence movement for Hong Kong, all in an attempt to undermine sympathy for the protestors, discredit their goals, and improve China's image internationally.

EA: Eliciting a specific emotional response

Disinformation was initially shared by state-supported media, such as CCTV or *China Daily,* before being re-shared on platforms such as Weibo, a popular Chinese blogging site, Twitter, and Facebook.[80] Posts relied on emotive language and imagery intended to simulate patriotic feelings, fear, or disgust, and avoided any substantive discussion of the issues at play.[81]

One particularly powerful approach relied on using emotionally arousing content to discredit and delegitimize the protestors. Video footage of violent protests was promoted on social media and labeled a precursor to terrorism, while coverage of the peaceful protests was censored. Any effort to contextualize the protests or express sympathy for the protestors was quickly deleted.[82]

Images, videos, and events were also manipulated or misrepresented. A female protestor lost her eye after being hit by a rubber bullet, likely shot by a police officer. In an effort to prevent a shift in sympathy, however, a widely shared video on Weibo allegedly showed the protestor accepting payment from other protestors, effectively insinuating the incident was staged.[83] During the campaign, the protestors were represented as violent thugs and the police as courageous heroes.[84] Many posts also identified the protestors with various types of insects, including cockroaches and locusts, intending to provoke disgust and perhaps imply that they, like insects, deserve extermination (Figure 17).[85]

[80] Diresta et al., *Telling China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives*, p. 22.

[81] Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 7.

[82] Krassi Twigg and Kerry Allen, "The disinformation tactics used by China," *BBC*, March 12, 2020, accessed https://www.bbc.com/news/56364952, July 20, 2021.

[83] Emily Feng, "China State Media Present Their Own Version of Hong Kong Protests," *NPR*, Aug. 14, 2019, accessed July 19, 2021, https://www.npr.org/2019/08/14/751039100/china-state-media-present-distorted-version-of-hong-kong-protests.

[84] Steven Meyers and Paul Mozur, "China is Waging a Disinformation War Against Hong Kong Protestors," *New York Times*, Aug. 13, 2019, accessed July 18, 2021, https://www.nytimes.com/2019/08/13/world/asia/hong-kong-protests-china.html.

[85] Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 4.

Figure 17. Tweet comparing protestors to cockroaches

Source: Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 7.

In taking this approach, China not only circulated emotionally arousing content, but effectively controlled the type of emotional arousal that the content would provoke. The circulating content simply didn't lend itself to an interpretation that might lead a viewer to feel sympathy, camaraderie, or pity for the protestors. It was, instead, specifically designed to elicit the emotions that China wanted people to feel. The disinformation intentionally: played up the violence of the protestors (fear), argued the protests were a precursor to terrorism (fear),

claimed the protests were funded by the West (surprise), and compared the protestors to cockroaches or other vermin (disgust).

CD: Eliminating cognitive dissonance

China's "Great Firewall" allows the CCP to control almost all media and social media content within mainland China. Countless websites are blocked, embedded censors delete unacceptable content, and people who speak out or share content can be arrested.[86] This control—in combination with the spread of disinformation—ensured that mainland Chinese were experiencing an almost entirely different series of events, described by some as a "parallel universe of narratives."[87] What was viewed by many in Hong Kong and around the world as a popular, largely peaceful demonstration movement was seen in China as a small, violent group of protestors that were unsupported by residents and provoked by foreign agents (Figure 18 and Figure 19).[88]

Figure 18. A tweet alleging that other actors were behind the protests

Source: Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 5.

[86] Meyers and Mozur, "China is Waging a Disinformation War Against Hong Kong Protestors."

[87] Meyers and Mozur, "China is Waging a Disinformation War Against Hong Kong Protestors."

[88] Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 2; Meyers and Mozur, "China is Waging a Disinformation War Against Hong Kong Protestors."

Figure 19. A state-sponsored news organization exaggerating protestors' violence

Source: Dotson, "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong," p. 7.

Critically, the PRC's ability to control the media ecosystem—both by sharing disinformation and by suppressing problematic truths—ensured that consumers were offered a narrative free of cognitive dissonance (i.e., a narrative of events with no internal contradictions). Those who remained within the Chinese media ecosystem simply weren't confronted with contradictory or dissonance-producing content. And because the Chinese media ecosystem eliminated the experience of cognitive dissonance, it also eliminated the need to engage in compensatory actions. As a result, viewers were never challenged to question the veracity of the dominant (though false) narrative being circulated. Spared the uncomfortable feeling of cognitive dissonance, they never had to change their behaviors, ignore contradictory data, compare narratives, or redouble their commitment to their original beliefs. This not only made the information appealing, but also increased the likelihood that people would believe the information: people are more likely to believe the narrative that creates the least dissonance.

#CoronaJihad

Primary mechanisms: GBN, EA

Background

Over the past 18 months, the spread of COVID-19 has been accompanied by the spread of disinformation about everything from the origins of the disease to potential cures. Although some circulating disinformation has been vague, some is part of a coordinated campaign. In mid-March 2020, a disinformation campaign took shape in India; this campaign implicated Indian Muslims in India's initial surge in cases through the use of memes, videos, tweets, and even statements from political actors blaming Muslims for intentionally spreading the virus. This campaign was not coordinated by a state actor—though the Indian government did participate at times—but was more of an organic grassroots effort that the government at times embraced and at times rejected.

The campaign can be traced to a gathering of 8,000 Tablighi Jamaat (an Islamic reformist group) preachers from over 40 countries, which was held at a mosque just outside of Delhi on March 13-15, 2020.[89] COVID-19 cases across India were just beginning to surge, and the first nationwide lockdown wasn't called until 10 days later, on March 25. The gathering consequently didn't violate any laws or restrictions; nor could it be considered a particularly irresponsible act or an outlier, because dozens of other large religious and political gatherings occurred around the same time. This particular gathering of Muslims, though, quickly became the scapegoat for India's surge in cases.[90] By March 29, reports emerged that there was a link between the Tablighi gathering and at least 24 new COVID-19 cases. And by March 31 the hashtag "coronajihad" was trending on Twitter.[91]

The Indian government, run by Narendra Modi and the pro-Hindu Bharatiya Janata Party (BJP), does not appear to have orchestrated this campaign. However, after initially downplaying the threat posed by COVID-19, they began to publicly blame India's surge in cases on the Tablighi gathering (Figure 20). The Health Secretary called the congregation out by name in daily briefings and some BJP politicians called the Tablighi gathering "coronaterrorism" and a

[89] Aniruddha Goshal, Sheikh Saaliq, and Emily Schmall, "Indian Muslims face stigma, blame for surge in infections," *AP News*, Apr. 25, 2020, accessed July 1, 2021, https://apnews.com/article/ad2e96f4caa55b817c3d8656bdb2fcbd.

[90] T. Soundararajan et al., *Coronajihad: An Analysis of COVID-19 Hate Speech and Disinformation*, Equality Labs, 2020, p. 2, accessed July 3, 2021, https://static1.squarespace.com/static/58347d04bebafbb1e66df84c/t/5ed86655611dc04dc4c48e7f/1591240284877/CORONAJIHAD_EqualityLabs_Report2020.

[91] Shweta Desai and Amarnath Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, Institute of Study of Diplomacy, 2020, p. 17, https://strongcitiesnetwork.org/en/wp-content/uploads/sites/5/2020/06/CoronaJihad.pdf.

"Talibani crime."[92] Some believe that the government highlighted the Tablighi gathering as the cause of the surge in order to redirect attention and hide the government's own mismanagement of the pandemic.[93]

Figure 20. A Republic TV panel demanding a crackdown on Tablighi Jamaat members

Source: Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 19.

GBN/EA: An imminent threat, feelings of disgust, and old enmities

Perhaps unsurprisingly, the narrative quickly shifted from the Tablighi Jamaat congregation being a hot spot for cases, to the narrative that *all* Muslims were deliberately spreading coronavirus.[94] Disinformation, in the form of decontextualized videos and memes, began spreading across social media, with many posts shared by far-right Hindu nationalists or circulated within Hindu Nationalist Facebook groups.[95] A Vimeo video of Muslims licking plates (originally part of a campaign to prevent food waste) was shared widely, with superimposed text accusing Muslims of trying to spread the virus.[96] Another video of a Sufi prayer tradition was mislabeled a video of Muslims intentionally sneezing on each other at the Tablighi

[92] Sanjana Rajgarhia, *Targeted Harassment: The Spread of #CoronaJihad*, 2020, accessed July 5, 2021, https://mediamanipulation.org/case-studies/targeted-harassment-spread-coronajihad.

[93] Goshal, Saaliq, and Schmall, "Indian Muslims face stigma, blame for surge in infections."

[94] Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 2.

[95] Soundararajan et al., *Coronajihad: An Analysis of COVID-19 Hate Speech and Disinformation*, p. 42.

[96] Rajgarhia, *Targeted Harassment: The Spread of #CoronaJihad*.

gathering, while yet another—old, out-of-context—video purportedly displayed a Muslim man spitting on cops to spread the virus. Images comparing Muslims to terrorists (Figure 21 and Figure 22) or venomous snakes were also shared widely. In each case, the content was designed to link an arousing feeling (disgust, anger, fear) with an imminent threat (i.e., GBN). This powerful combination thus increased the likelihood that viewers would engage with, accept, and share the content.

Figure 21. An example of a video claiming Muslims are spreading the virus

Source: Rajgarhia, *Targeted Harassment: The Spread of #CoronaJihad*.

Figure 22. A caricature of a Tablighi member as a suicide bomber, with COVID-19 as explosives.

Source: Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 8.

Moreover, the rhetoric broadly fell into four categories—depicting Muslims as contaminated, uncivilized, deceptive, or anti-national jihadists—and built on preexisting Hindu-Muslim tensions in India (i.e., GBN).[97] Protests over a new citizenship law that excluded Muslims had been ongoing since December 2019, and escalated into deadly riots in February 2020. Thus, tensions between the Hindu and Muslim communities were already high before the Tablighi gathering and the COVID-19 pandemic. Islamophobic content had also been widespread on both Twitter and Facebook prior to the pandemic.[98]

The videos and memes spread rapidly. #Coronajihad appeared in almost 300,000 tweets between March 28 and April 3, which saw 700,000 points of engagement and which were seen by over 170 million users around the world.[99] In addition to Twitter, this content spread on

[97] Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 9.

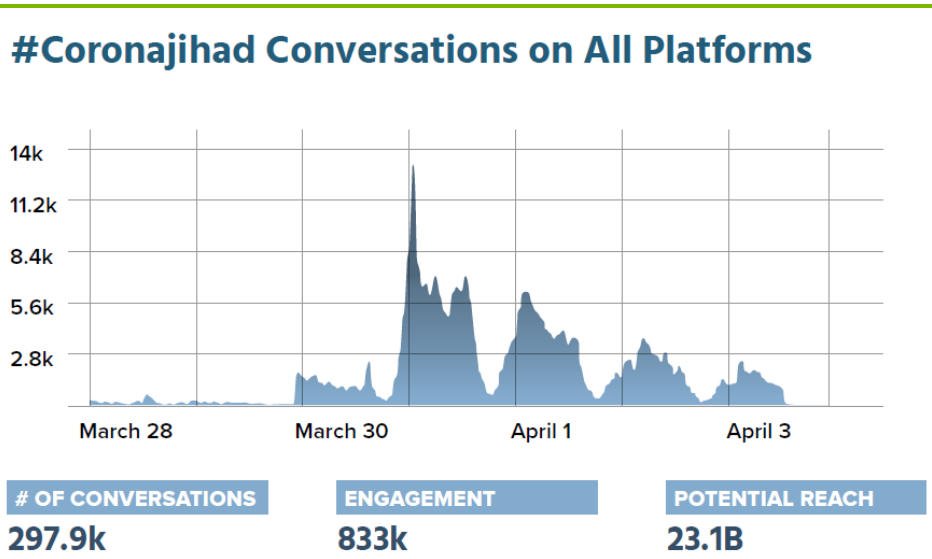
[98] Soundararajan et al., *Coronajihad: An Analysis of COVID-19 Hate Speech and Disinformation*, p. 20.

[99] Rajgarhia, *Targeted Harassment: The Spread of #CoronaJihad*.

Facebook, Vimeo, TikTok, YouTube, and WhatsApp, the top source of COVID-19 disinformation in India during the pandemic.[100]

Again, this particular disinformation campaign was not a coordinated effort crafted and executed by professional operators; instead, it was overwhelmingly driven by regular citizens who were participating organically, and (likely) sharing information within their own networks (i.e., GBN). BJP politicians' acceptance of (or willingness to exploit) the narrative amplified the effort by generating media exposure, with domestic media ultimately advancing the same, incorrect narrative about the role of Muslims in the spread of the virus (Figure 20).[101] India represents the largest market for Facebook, WhatsApp, Instagram, and YouTube, which meant that this disinformation could quickly make its way into the global discourse, where Islamaphobia helped the trend gain ground with users in other parts of the world. Between March 28 and April 3, the potential global reach of #coronajihad across all social media sites (except Facebook) and mainstream media was 23.1 billion users (Figure 23).

Figure 23. Tracking the reach of #Coronajihad



Source: Soundararajan et al., *Coronajihad: An Analysis of COVID-19 Hate Speech and Disinformation*, p. 18.

Although it is often difficult to track the direct effect of a disinformation campaign, many experts believe that this disinformation has led to discrimination, economic boycotts, and even violence against Muslims. One source notes that mob violence led to 200 injuries and 53 deaths, while another cites an Indian Muslim man's suicide as a result of being shunned by his

[100] Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 7.

[101] Rajgarhia, *Targeted Harassment: The Spread of #CoronaJihad*.

community.[102] It was only after the intervention of the WHO, UN, and broader international community that Modi tweeted in mid-April about the importance of national unity in the face of the virus and the government started to take actions to counter the disinformation.[103]

At the outset of this report, we noted that the distinction between disinformation and misinformation was not particularly relevant for a discussion that foregrounded the psychological mechanisms at play. Because the consumer rarely knows the intentions of a meme's original creator, it doesn't matter whether the content was unintentionally or intentionally inaccurate. This case highlights the reality that a disinformation campaign can also be effective without intentional cultivation. The #Coronajihad campaign was an organic, grassroots effort and it is highly unlikely that any of the individual content creators intentionally and knowingly aspired to hijack the psychological mechanisms of the consumers. This was likely done intuitively, but the effect is the same: content was created that elicited a range of intense emotions that spoke to an imminent threat, and that amplified pre-existing group conflicts (i.e., EA and GBN).

Summary

These five real-world examples of disinformation clearly illustrate the human vulnerability to what we might call psychological hacking. Each of these cases studies demonstrates how the manipulation of normal and functional psychological mechanisms can increase our receptivity to disinformation, and how those who create disinformation can leverage this vulnerability by designing content that is more likely to evade scrutiny or increase engagement.

The cases also illustrate some potentially real—and not merely notional—threats that the US faces, by describing disinformation campaigns that have already been executed. In doing so, they provide detail on the capabilities that our foreign adversaries already possess, and on the threat posed by grassroots campaigns that go unchecked.

This vulnerability cannot be fully mitigated. The psychological mechanisms outlined above—and discussed in more detail in our companion report—*The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*—are necessary for humans to function effectively.[104] Responding to these threats is critical.

---

[102] Rajgarhia, *Targeted Harassment*; Goshal, Saaliq, and Schmall, "Indian Muslims face stigma, blame for surge in infections."

[103] Desai and Amarasingam, *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India*, p. 17.

[104] Wolters et al., *Psychology of (Dis)information*.

Recommendations and Conclusions

The psychological mechanisms discussed above—and linked to the five case studies that this report explores—are not only normal, but critical to our ability to function. Moreover, the four mechanisms discussed in this report are a subset of a larger suite of mechanisms that are critical to human functioning. Unfortunately, necessary reliance on these mechanisms leaves us vulnerable to psychological hacking. Some of this hacking might be well intentioned (designed, as one example, to encourage smart choices about smoking), and some will be malicious (such as those explored in this report).[105]

Vectors for intervention

What can the US government (USG), and DOD as a major component, do about this situation? They can respond to the challenge posed by malicious disinformation—particularly that spread by foreign state and nonstate actors—through three broad vectors:

- At the source, by identifying and stopping producers of disinformation
- In the environment, through regulatory intervention, which will require legislative action
- With the audience, by reducing individual vulnerability to disinformation

Stopping the source

Though the efforts of organizations such as the Global Engagement Center and the Joint MISO WebOps Center, the USG is already actively working to disrupt disinformation at the source. The reality, though, is that stopping disinformation at the source is effectively a game of whack-a-mole that greatly favors the mole. Creating psychologically effective disinformation requires relatively minimal resources, and spreading this content can be facilitated by the use of increasingly low cost botnets that are available to rent.[106]

Addressing the environment

Addressing the problem of disinformation in the environment (i.e., on social media platforms) is largely outside the direct control of the USG, and certainly outside that of DOD. To date,

[105] L. Festinger, *Theory of Cognitive Dissonance*.; Stricklin and McBride, *Social Media Bots: Laws, Regulations, and Platform Policies*.

[106] McBride et al., *Social Media Bots: Implications for Special Operations Forces.*

technology has advanced faster than laws and regulations and it will likely continue to do so. This has left social media companies largely unregulated and means that buy-in from the companies themselves will be needed in order to bring about the changes that could reduce the threat posed by disinformation. Buy-in is critical, in part because disinformation is often sensational, and sensational content is key to the business models of social platforms. Clicks, likes, and views generate revenue for social media companies; as a result, the companies benefit from content that is sensational and arousing, and are not incentivized to decrease its circulation.

Social media companies are not powerless. They could require users to verify the veracity of a statement before posting; they could share video tips about detecting disinformation; and they could create banner headlines to share accurate news with users. Social media companies currently use algorithms to identify users' interests and share suggested content and target advertisements. There might be a way to leverage these, or other, algorithms to identify and flag disinformation or to allow only authentic users to share content. Again, the challenge is not in identifying the way, but in rallying the will.

Traditional media also have a role to play, especially since good journalism can lead to civic resilience to disinformation. If people know and trust reputable sources, they are less likely to be persuaded by non-reputable sources. To do this, however, reputable media would need to be less sensational and focus on providing the facts without spin, exaggeration, or bias. Similar to social media companies, traditional media would need to work against the incentives of speed and sensation (breaking news banners and sensational headlines drive engagement), which are also critical to the financial health of traditional media companies.

Protecting the audience

When it comes to the potential audiences for disinformation, the USG again is left with little that it can do directly to protect the totality of its citizens. Within the freedom-loving culture of the United States, Americans are not going to accept the government telling them what information they can and cannot—or even what they should or should not—access via social or traditional media platforms. That said, there is a sizeable subset of US citizens over which the USG does have a line of direct control: USG employees, including all of DOD (service members and civilians). And there are things the USG/DOD could do to help protect those employees, both from the direct effects of disinformation and from serving as unwitting vectors of disinformation.

To do this, it would be critical for the USG/DOD to invest in the development and deployment of non-partisan, evidence-based interventions that will protect employees (and uniformed service members as a key component) against this content. As we have seen in this report, disinformation has already been used to influence perceptions of allies and adversaries;

increase distrust of neighbors and friends; distort understanding of current events; affirm that unflattering stereotypes are valid; and leverage worries and fears to prompt action. Deployed expertly, disinformation can influence US partnerships, increase domestic discord and violence, cripple the US economy, and endanger the health and well-being of the US population. Thus, by protecting the health of USG/DOD employees—in this case, protecting them from psychological manipulation—the USG protects those tasked to protect and serve the country, and protects the country from government members who might otherwise unwittingly be vectors of foreign adversary disinformation.

Protecting the USG

Actions to protect USG employees and servicemembers both from the threat posed by disinformation itself, and from being unwitting vectors for harming the nation through its spread, have the potential to be quite powerful. They could undertake actions such as designing and implementing a non-partisan, evidence-based set of interventions to protect against the innate human vulnerabilities to disinformation that were described above. The goal of countering disinformation, after all, is not to shape opinion, but to prevent opinion from being shaped by inaccurate data.

DOD, for example, has long recognized its duty and responsibility both to protect those that serve the nation, and to protect the nation from adversaries that would use these servicemembers as vectors for harming our country. Additionally, DOD has already identified force protection as a core element of its approach to disinformation. In a March 2021 statement to the House Armed Services Committee Subcommittee on Intelligence and Special Operations, DOD represented noted the following:

> Our Soldiers, Sailors, Marines, Airmen, Guardians, civilians, and their families are part of the American public directly targeted by malign actors' disinformation, misinformation, and propaganda. DoD views this as a critical force protection issue. The Services are proactively leading efforts to enable resilience against these threats. Enabling the force to recognize deceptive information tactics by adversarial information operations, developing digital literacy, and employing critical thinking skills are a few key initiatives within this line of effort.[107]

DOD already conducts interventions on other information threat vectors by providing servicemembers with regular cybersecurity training designed to ensure that they do not

[107] Christopher Maier, Neill Tipton and James Sullivan, Joint Statement for the Record before the House Armed Services Committee, Subcommittee on Intelligence and Special Operations, "Disinformation in the Gray Zone: Opportunities, Limitations, Challenges," March 16, 2021, https://docs.house.gov/meetings/AS/AS26/20210316/111323/HHRG-117-AS26-Wstate-MaierC-20210316.pdf.

unwittingly help foreign state and nonstate actors penetrate USG computer networks. Disinformation training would work similarly, in that it would alert USG employees and servicemembers to the threat posed by disinformation, and train them to be alert to its influence.

Evidence-based interventions

Designing and implementing a series of interventions that would reduce the influence and effect of disinformation is thus a course of action that aligns with DOD's stated priorities in this space. Below we present a preliminary list of recommendations.

Preventative inoculation

- **Game-based inoculation:** Some research has shown the potential benefits of using game-based inoculation focused on disinformation, with examples including "Go Viral" and "Bad News." This game-based training is not dissimilar to a viral vaccine and works best when a "dose" is given before encountering disinformation and "boosters" (in the form of additional training) are provided to sustain the resistance.
- **Awareness campaigns:** Awareness campaigns—whether videos, public service announcements, or advertisements on social media platforms themselves—could help increase individuals' knowledge of the issue of disinformation and make individuals more wary and discerning when consuming media online.
- **Video and banner reminders:** There is evidence that these types of smaller interventions, whether 60-to-90-second videos or banner reminders sharing tips on identifying disinformation, can prime users to be more aware of the issue.

Cultivation of deeper, analytic thinking

- **Media literacy training:** Media literacy training can take different forms and, in some instances, might even include game-based inoculation. The goal of media literacy training is to strengthen individuals' ability to assess and critically evaluate media. Training and education can help people to recognize the emotions, biases, and stereotypes present when they consume media. Training could focus on increasing critical thinking, source evaluation, and emotional manipulation.
- **Confirmation before posting:** Asking an individual to verify the veracity of a claim, tweet, or post before they share it can slow their sharing of information. Some social media companies already do this. For example, Twitter now asks individuals if they're sure they want to retweet a post if they haven't clicked on the link and read the article. These type of pop-up windows prime individuals to think through the source and the content, and confirm their intent to share.

- **Pop-up windows alerting individuals to unregulated sites:** Although it is impossible to stop users from traveling to sites with disinformation present, pop-up windows could alert users to sites ran by bad actors or unregulated social media sites.
- **Pop-up windows sharing accurate information:** During the COVID-19 pandemic some social media companies, including Facebook and WhatsApp, developed features to directly share accurate information with users who encountered disinformation. Social media posts that mentioned COVID-19 vaccines, for example, would also be accompanied by a banner sharing accurate information on the issue and additional reliable sources. These features could be implemented more broadly across platforms.
- **Fact checking and verification:** Sites that provide fact checking, explainers, and analysis help to debunk claims and expose fake news, providing individuals with accurate information on contentious issues. There are some examples of promising tools that allow the public to do things such as check the authenticity of an image.

Conclusion

The threat posed by disinformation is significant; according to the DOD, it "poses one of today's greatest challenges to the United States, not just to DoD."[108] Responding to this challenge will require a whole-of-government response that stops disinformation at its source, slows the flow of information through the digital environment, and decreases our vulnerability to this type of content. Such a response will take years to develop and implement, but there are immediate actions that the DOD can take on a much shorter timeline.

Specifically, the DOD could implement IT changes that would allow pop-up banners indicating unverified information is present and requests the user to confirm they want to share information. Both of these actions would prompt the DOD IT user to think more deeply about the information before they share it. DOD can use the results of this literature review to inform-and sponsor-the research necessary to design a non-partisan, evidence-based intervention to protect employees and servicemembers from being manipulated by disinformation through both preventative inoculation and cultivating deeper, analytical thinking approaches. This type of intervention is well within the norms and traditions of the government's previous actions to protect both its employees, service members, and, by extension, our nation. Further, there are already commercially available games that may

[108] Maier, Tipton, and Sullivan, "Disinformation in the Gray Zone: Opportunities, Limitations, Challenges."

address this training gap (e.g., Go Viral[109]). As we have highlighted in this report, disinformation is something that any consumer of online media is likely to encounter, and our normal and adaptive psychological mechanisms enable it to act on our brains in ways that can be beneficial or malicious. Although it is beyond the scope of the USG and DOD's roles and responsibilities to directly protect all Americans from the effect of disinformation, it can do more to prevent employees and servicemembers from it. Such additional actions are critical, given the threat posed by disinformation and the inadequacy of existing means to stop its influence.

[109] Based on discussion with SME, Apr. 2021.

Figures

Tables

References

Aleksejeva, Nika, Lukas Andriukaitis, Luiza Bandeira, Donara Barojan, Graham Brookie, Eto Buziashvili, Andy Carvin, et al. *Operation "Secondary Infektion": A Suspected Russian Intelligence Operation Targeting Europe and the United States.* Digital Forensic Research Laboratory, 2019. Accessed June 8, 2021. https://www.atlanticcouncil.org/in-depth-research-reports/report/operation-secondary-infektion/.

Asser, Martin. "What the Muhammad Cartoons Portray." BBC News. Jan. 2, 2010. http://news.bbc.co.uk/2/hi/middle_east/4693292.stm.

Berger, Jonah. "Arousal Increases Social Transmission of Information." *Psychological Science* 22, no. 6 (2011): 891-893. https://journals.sagepub.com/doi/10.1177/0956797611413294 .

BBC. "Russian trolls' chief target was 'black US voters' in 2016." *BBC*. Oct. 9, 2019. Accessed July 16, 2021. https://www.bbc.com/news/technology-49987657.

Brooking, Emerson T., and Suzanne Kianpour. *Iranian Digital Influence Efforts.* Atlantic Council. 2020. https://www.atlanticcouncil.org/wp-content/uploads/2020/02/IRAN-DIGITAL.pdf.

———. *Iran Digital Influence Efforts: Guerrilla Broadcasting for the Twenty-First Century.* Atlantic Council. 2020. Accessed July 6, 2021. https://www.atlanticcouncil.org/in-depth-research-reports/report/iranian-digital-influence-efforts-guerrilla-broadcasting-for-the-twenty-first-century/.

Brooks, B.P., N. DiFonzo, and D.S. Ross. "The GBN-dialogue model of outgroup-negative rumor transmission: Group Membership, belief, and novelty." *Nonlinear Dynamics, Psychology, and Life Sciences* 17, no. 2 (2013): 269-293.

Carnegie, Dale. *How to win friends and influence people*. Simon & Schuster, 1936.

Center for Internet Security. "MS-ISAC Security Primer: Typosquatting." Feb. 2018. Accessed July 14, 2021. https://www.cisecurity.org/wp-content/uploads/2018/02/MS-ISAC-Typosquatting-Security-Primer.pdf.

Cialdini, Robert B. *Influence: Science and Practice*. 4th ed. Boston: Allyn and Bacon, 2001.

Cook, Sarah. "Welcome to the New Era of Chinese Government Disinformation." *The Diplomat.* May 11, 2020. https://thediplomat.com/2020/05/welcome-to-the-new-era-of-chinese-government-disinformation/.

Counter Extremism Project. "IRGC (Islamic Revolutionary Guard Corps)." Accessed Feb. 3, 2021. https://www.counterextremism.com/threat/irgc-islamic-revolutionary-guard-corps.

Deibert, Ronald. "Endless Mayfly: An Invasive Species in the Social Media Ecosystem." The Citizen Lab. May 14, 2019. Accessed July 10, 2021. https://deibert.citizenlab.ca/2019/05/endless-mayfly/.

Desai, Shweta, and Amarnath Amarasingam. *CoronaJihad: COVID-19, Misinformation and Anti-Muslim Violence in India.* Institute of Study of Diplomacy. 2020. https://strongcitiesnetwork.org/en/wp-content/uploads/sites/5/2020/06/CoronaJihad.pdf.

DiResta, Renee, Carly Miller, Vanessa Molter, John Pomfret, and Glenn Tiffert. *Telling China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives.* Stanford Internet Observatory. 2020. Accessed July 18, 2021. https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/sio-china_story_white_paper-final.pdf.

DiResta, Renee, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, and Ben Johnson. *The Tactics & Tropes of the Internet Research Agency.* New Knowledge. DigitalCommons@University of Nebrasksa - Lincoln. 2019. Accessed July 1, 2021. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1003&context=senatedocs.

Dotson, John. "Chinese Covert Social Media Propaganda and Disinformation Related to Hong Kong." *China Brief* 19, no. 16 (2019): 1-7. Accessed July 20, 2021. https://jamestown.org/wp-content/uploads/2019/09/Read-the-09-06-2019-CB-Issue-in-PDF.pdf

Feng, Emily. "China State Media Present Their Own Version of Hong Kong Protests." *NPR*. Aug. 14, 2019. Accessed July 19, 2021. https://www.npr.org/2019/08/14/751039100/china-state-media-present-distorted-version-of-hong-kong-protests.

Festinger, L. *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson, 1957.

FireEye. "Suspected Iranian Influence Operations: Leveraging Inauthentic News Sites and Social Media Aimed at U.S., U.K, Other Audiences." FireEye. Aug. 21, 2018. Accessed July 20, 2021. https://www.fireeye.com/blog/threat-research/2018/08/suspected-iranian-influence-operation.html.

Fischer, Sara. ""Unreliable" news sources got more traction in 2020." Axios. Dec. 22, 2020. https://www.axios.com/unreliable-news-sources-social-media-engagement-297bf046-c1b0-4e69-9875-05443b1dca73.html.

Fiske, S.T., and S.E. Taylor. *Social Cognition: From Brains to Culture*. United Kingdom: SAGE Publications, 2016.

Gerasimov, V. V. "The Value of Science is Foresight." Ценность науки в предвидении. VPK. ВПК. Feb. 26, 2013. https://vpk-news.ru/articles/14632.

Ghosh, Dipayan. "Are We Entering a New Era of Social Media Regulation?" *Harvard Business Review*. Jan. 14, 2021. Accessed Aug. 19, 2021. https://hbr.org/2021/01/are-we-entering-a-new-era-of-social-media-regulation.

Goshal, Aniruddha, Sheikh Saaliq, and Emily Schmall. "Indian Muslims face stigma, blame for surge in infections." *AP News*. Apr. 25, 2020. Accessed July 1, 2021. https://apnews.com/article/ad2e96f4caa55b817c3d8656bdb2fcbd.

Harding, Luke. "Russians 'Spread Fake Plot to Assassinate Boris Johnson' on Social Media." *The Guardian*. June 22, 2019. Accessed July 12, 2021. https://www.theguardian.com/world/2019/jun/22/russians-spread-fake-plot-to-assassinate-boris-johnson.

Hopkins, Tatyana. "Social media companies profiting from misinformation." GWToday. June 19, 2020. Accessed Sept. 27, 2021. https://gwtoday.gwu.edu/social-media-companies-profiting-misinformation.

Howard, Philip N., Bharath Ganesh, Dimitra Liotsiou, John Kelly, and Camille Francois. *The IRA, Social Media and Political Polarization in the United States, 2012-2018.* Computational Propaganda Research Project. 2019. Accessed July 5, 2021. DigitalCommons@University of Nebraska - Lincoln. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1004&context=senatedocs.

Huguet, Alice, Jennifer Kavanagh, Garrett Baker, and Marjory S. Blumenthal. *Exploring Media Literacy Education as a Tool for Mitigating Truth Decay.* RAND. RR-3050-RC. 2019. doi: https://doi.org/10.7249/RR3050.

Jackson, Dean. "Issue Brief: How Disinformation Impacts Politics and Publics." National Endowment for Democracy. May 29, 2018. https://www.ned.org/issue-brief-how-disinformation-impacts-politics-and-publics/.

Kang, Ceclia. "Lawmakers, Taking Aim at Big Tech, Push Sweeping Overhaul of Antitrust." *The New York Times*. June 11, 2021. Accessed Aug.19, 2021. https://www.nytimes.com/2021/06/11/technology/big-tech-antitrust-bills.html.

Keating, Dan, Kevin Schaul, and Leslie Shapiro. "The Facebook ads Russians targeted at different groups " *Washington Post*. Nov. 1, 2017. https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/.

Kim, Young Mie. *New Evidence Shows How Russia's Election Interference Has Gotten More Brazen.* 2020. Accessed July 7, 2021. https://www.brennancenter.org/our-work/analysis-opinion/new-evidence-shows-how-russias-election-interference-has-gotten-more.

———. *Uncover: Strategies and Tactics of Russian Interference in US Elections.* Project DATA. 2018. Accessed July 8, 2021. https://journalism.wisc.edu/wp-content/blogs.dir/41/files/2018/09/Uncover.Kim_.v.5.0905181.pdf.

Kliman, Daniel, Andrea Kendall-Taylor, Kristine Lee, Joshua Fitt, and Carisa Nietsche. *Dangerous Synergies.* CNAS. 2020. https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/CNAS-Report-Dangerous-Synergies-May-2020-DoS-Proof.pdf?mtime=20200506164642&focal=none.

Kurlantzick, Joshua. "How China Ramped Up Disinformation Efforts During the Pandemic." Council on Foreign Relations. Sept. 10, 2020. https://www.cfr.org/in-brief/how-china-ramped-disinformation-efforts-during-pandemic.

Lee, Doowan. "The United States Isn't Doomed to Lose the Information Wars." *Foreign Policy.* Oct. 16, 2020. https://foreignpolicy.com/2020/10/16/us-election-interference-disinformation-china-russia-information-warfare/.

Lim, Gabrielle, Etienne Maynier, John Scott-Railton, Alberto Fittarelli, Ned Moran, and Ron Deibert. *Burned After Reading: Endless Mayfly's Ephemeral Disinformation Campaign.* 2019. Accessed July 9, 2021. https://citizenlab.ca/2019/05/burned-after-reading-endless-mayflys-ephemeral-disinformation-campaign.

Lockhart, P. R. "How Russia exploited racial tensions in America during the 2016 elections." *Vox*. Dec. 17, 2018. Accessed July 7, 2021. https://www.vox.com/identities/2018/12/17/18145075/russia-facebook-twitter-internet-research-agency-race

Lord, Kristin M., and Katya Vogt. "Strengthen Media Literacy to Win the Fight Agaisnt Misinformation." *Stanford Social Innovation Review* (2021). Accessed Aug. 18, 2021. https://ssir.org/articles/entry/strengthen_media_literacy_to_win_the_fight_against_misinformation.

Maier, Christopher, Neill Tipton, and James Sullivan. "Joint Statement for the Record before the House Armed Services Committee, Subcommittee on Intelligence and Special Operations on 'Disinformation in the Gray Zone: Opportunities, Limitations, Challenges.'" Mar. 16, 2021. https://docs.house.gov/meetings/AS/AS26/20210316/111323/HHRG-117-AS26-Wstate-MaierC-20210316.pdf.

McBride, Megan, Zack Gold, Jonathan Schroden, and Lauren Frey. *Cryptocurrency: Implications for Special Operations Forces.* CNA. 2019. Accessed Aug. 17, 2021. https://shb2016docweb.cna.org:8443/dctmsearch/FFRDC/Publications/2019/CRM-2019-U-020186/CRM-2019-U-020186-Final.pdf?contentTicket=1lf662rjvmnqi1b780to&Reload=1629243048888&__dmfClientId=1629243040318.

McBride, Megan, Zack Gold, and Kasey Stricklin. *Social Media Bots: Implications for Special Operations Forces.* CNA. 2020. Accessed Aug 17, 2021. https://www.cna.org/CNA_files/PDF/DRM-2020-U-028199-Final.pdf.

Meyers, Steven, and Paul Mozur. "China is Waging a Disinformation War Against Hong Kong Protestors." *New York Times*. Aug. 13, 2019. Accessed July 18, 2021. https://www.nytimes.com/2019/08/13/world/asia/hong-kong-protests-china.html.

Michel, Casey. *Website targeting black Americans appears to be elaborate Russian propaganda effort.* ThinkProgress. 2017. Accessed July 20, 2021. https://thinkprogress.org/black-matters-us-site-90625b18f262/.

Mikkelsen, Barbara. "The Killer in the Backseat". Urban Legends Reference Pages. http://www.snopes.com/horrors/madmen/backseat.htm.

Nimmo, Ben, Camille François, C. Shawn Eib, Lea Ronzaud, Rodrigo Ferreira, Chris Hernon, and Tim Kostelancik. *Exposing Secondary Infektion.* Graphika. 2020. Accessed July 9, 2021. https://secondaryinfektion.org/.

Rajgarhia, Sanjana. *Targeted Harassment: The Spread of #CoronaJihad.* 2020. Accessed July 5, 2021. https://mediamanipulation.org/case-studies/targeted-harassment-spread-coronajihad.

Rawnsley, Adam. "Russian Trolls Hype Coronavirus and Giuliani Conspiracies." *Daily Beast*. Apr. 9, 2020. Accessed July 13, 2021. https://www.thedailybeast.com/russian-trolls-hype-coronavirus-and-giuliani-conspiracies.

Reddit. "IRA enlists Muslim militants!" Reddit Post by u/robeharty. 2019. https://www.reddit.com/r/ireland/comments/bfzfbb/ira_enlists_muslim_militants/.

Robinson, Martin. "Russia Has Tried to Reignite the Troubles With Fake Social Media Posts." *Daily Mail*. June 26, 2019. Accessed July 14, 2021. https://www.dailymail.co.uk/news/article-7182739/Russia-tried-reignite-Troubles-fake-social-media-posts.html.

Shane, Scott, and Sheera Frenkel. "Russian 2016 Influence Operation Targeted African-Americans on Social Media." *New York Times*. Dec. 17, 2018. Accessed July 11, 2021. https://www.nytimes.com/2018/12/17/us/politics/russia-2016-influence-campaign.html.

Silverman, Craig, and Jane Lytvynenko. Screenshot of lookalike *Guardian* page via "How A Hoax Made To Look Like A Guardian Article Made Its Way To Russian Media." Buzzfeed News. Aug. 15, 2017. Accessed Aug. 6, 2021. https://www.buzzfeednews.com/article/craigsilverman/how-a-hoax-made-to-look-like-a-guardian-article-made-its.

Soundararajan, T., A. Kumar, P. Nair, and J. Greely. *Coronajihad: An Analysis of COVID-19 Hate Speech and Disinformation.* Equality Labs. 2020. Accessed July 3, 2021. https://static1.squarespace.com/static/58347d04bebafbb1e66df84c/t/5ed86655611dc04dc4c48e7f/1591240284877/CORONAJIHAD_EqualityLabs_Report2020.

Stricklin, Kasey. "Why Does Russia Use Disinformation?" Lawfare. Mar. 29, 2020. https://www.lawfareblog.com/why-does-russia-use-disinformation.

Stricklin, Kasey, and Megan K. McBride. *Social Media Bots: Laws, Regulations, and Platform Policies.* CNA. 2020. Accessed Aug. 17, 2021. https://shb2016docweb.cna.org:8443/dctmsearch/FFRDC/Publications/2020/DIM-2020-U-028193/DIM-2020-U-028193-Final.pdf?contentTicket=1q2dvk32st5blr1b732jk&Reload=1629242886772&__dmfClientId=1629242878135.

Twigg, Krassi, and Kerry Allen. "The disinformation tactics used by China." *BBC*. Mar. 12, 2020. Accessed July 20, 2021. https://www.bbc.com/news/56364952.

US House of Representatives Permanent Select Committee on Intelligence. "Exposing Russia's Effort to Sow Discord Online: The Internet Research Agency and Advertisements." Accessed Sept. 27, 2021. https://intelligence.house.gov/social-media-content/.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The spread of true and false news online." *Science* 359, no. 6380 (2018): 1146-1151. https://science.sciencemag.org/content/359/6380/1146.

West, Darrell M. *How to Combat Fake News and Disinformation.* Brookings Institution. 2017. https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation/.

Wolters, Heather, Kasey Stricklin, Neil Carey, and Megan K. McBride. *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms.* CNA. DRM-2021-U-029337-1Rev. Sept. 2021.

Zakem, Vera, Megan K. McBride, and Kate Hammerberg. *Exploring the Utility of Memes for U.S. Government Influence Campaigns.* CNA. 2018. Accessed Aug. 17, 2021. https://shb2016docweb.cna.org:8443/dctmsearch/FFRDC/Publications/2018/DRM-2018-U-017433/DRM-2018-U-017433-Final.pdf?contentTicket=19gr23r22u4f1j1b6q4pv&Reload=1629242594111&__dmfClientId=1629242585249.

This report was written by CNA's Strategy, Policy, Plans, and Programs Division (SP3).

SP3 provides strategic and political-military analysis informed by regional expertise to support operational and policy-level decision-makers across the Department of the Navy, the Office of the Secretary of Defense, the unified combatant commands, the intelligence community, and domestic agencies. The division leverages social science research methods, field research, regional expertise, primary language skills, Track 1.5 partnerships, and policy and operational experience to support senior decision-makers.

Dedicated to the Safety and Security of the Nation

DRM-2021-U-030881-Final

3003 Washington Boulevard, Arlington, VA 22201
www.cna.org ● 703-824-2000